SEP 18 1930

PERIODICAL ROOM
GENERAL LIBRARY
UNIV. OF MICH.

The Journal of EDUCATIONAL PSYCHOLOGY

Devoted Primarily to the Scientific Study of Problems of Learning and Teaching.

HAROLD RUGG, *Editor*

VOL. XXI, No. 6 SEPTEMBER, 1930 $5.00 per year

CONTENTS

INDIVIDUAL DIFFERENCES IN FORGETTING PLEASANT AND UNPLEASANT EXPERIENCES. *H. Meltzer* 399

THE INFLUENCE OF ATTITUDE UPON TRANSFER. *Mattie F. Dorsey and L. Thomas Hopkins* 410

HABIT AND ATTITUDE. *Harold Saxe Tuttle* 418

THE RELATION OF NERVOUS-EMOTIONAL STABILITY TO EDUCATIONAL ACHIEVEMENT. *Noel Keys and G. H. Whiteside* . . . 429

THE BIOLOGICAL SIGNIFICANCE OF INTELLIGENCE TESTS. *David W. Oates* 442

ADULT AGES AS A FACTOR IN LEARNING. *Herbert Sorenson* . . . 451

THE VALIDITY OF TEACHERS' JUDGMENTS OF DIFFICULTY IN CURRICULAR MATERIAL. *H. E. Smith* 460

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION 467

PUBLICATIONS RECEIVED 473

Published Monthly Except June to August by

Warwick & York, Inc.,

10 East Centre Street Baltimore, Md.

Entered as Second Class matter November 15, 1921, at the Post Office at Baltimore, Md., under the Act of March 3, 1879; additional entry as second class matter at York, Pa.

THE JOURNAL OF EDUCATIONAL PSYCHOLOGY

ESTABLISHED 1910

Editor: HAROLD RUGG

Board of Associate Editors

J. CARLETON BELL — ARTHUR I. GATES — RUDOLF PINTNER — E. L. THORNDIKE
FRANK N. FREEMAN — V. A. C. HENMON — L. M. TERMAN — GOODWIN B. WATSON

Assistant to the Editor: FRANCES M. FOSTER

The Journal is published monthly, except June to August, by Warwick and York, Inc., Baltimore, Md., and York, Pa.

The price is $5.00 a year, payable in advance; foreign postage 30 cents extra. Single current issues are 70 cents each, and subscriptions for less than a year are charged at single issue rate for each number ordered. Back volumes are $6.00 each; single back numbers, 90 cents each, except, when more than three years old, the price is $1.00 each.

Subscribers should notify the publishers of change in address at least four weeks before publication of the issue with which the change is to take effect. Claims for non-receipt of an issue will not be honored unless made within two weeks after receipt of the next succeeding number.

Manuscripts for publication, books or other materials for review, and correspondence concerning editorial matters should be addressed to Harold Rugg, Editor, 425 West 123rd Street, New York City. Subscriptions should be sent to the publishers.

WARWICK AND YORK, INC., *Baltimore, Maryland*

TEACHERS' HEADQUARTERS

European

Plan

Entirely

Fire-Proof

Centrally

Located

EDWARD DAVIS

Manager

HOTEL RENNERT

BALTIMORE, MARYLAND

SCHOOL AND SOCIETY

A weekly journal covering the field of education in relation to the problems of American democracy.

The objects of SCHOOL AND SOCIETY are the advancement of education as a science and the adjustment of our lower and higher schools to the needs of modern life.

It encourages discussion of all problems from every point of view, it pays special attention to foreign affairs; it publishes accounts of educational events of the past week and research on fundamental problems.

Subscription $5.00—Single Numbers 15 Cents

[*SCHOOL AND SOCIETY will be sent to new subscribers for thirteen weeks for one dollar, which may be sent at the risk of the publisher.*]

THE SCIENCE PRESS

Grand Central Terminal - - New York, N. Y.

INTRODUCTORY PSYCHOLOGY FOR TEACHERS

By EDWARD K. STRONG, JR.

$2.20 + 12¢ postage

WARWICK & YORK, INC., *Baltimore, Maryland*

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume XXI | September, 1930 | Number 6

INDIVIDUAL DIFFERENCES IN FORGETTING PLEASANT AND UNPLEASANT EXPERIENCES

H. MELTZER

Psychiatric Clinic of St. Louis and Washington University

The rôle of feeling in human remembering and forgetting is an important problem in both affective psychology[1] and the psychology of learning.[2] Representing these two interests we have two explanatory principles concerning the relationship of feeling and learning. These are respectively the Freudian doctrine of repression and Thorndike's "law of effect." The interest of psychoanalysts is in mechanisms underlying dissociation and amnesia; that of educational psychologists is in mechanisms of the learning process. Expressed in psychoanalytic verbiage, the aspect of Freudian doctrine relevant for our present purpose says, in effect, that experiences charged with unpleasant feeling tone are relegated to the unconscious and kept there by the censor.

According to the law of effect, it is the pleasant (satisfying) or unpleasant (annoying) consequences of an act which determines the repetition of that act. That unpleasant experiences are more easily forgotten than pleasant is implied in both psychoanalysis and educational psychology—more directly, however, in the former. But implications are not facts. What are the facts concerning individual differences in human experiencing of pleasant and unpleasant activities? Concerning individual differences in human remembering and forgetting of pleasant and unpleasant personal experiences?

[1] In his contribution to "Foundations of Experimental Psychology," called "The Expressions of Emotions" (Chap. XIII), Landis speaks as follows of the larger problem of which this is one aspect: "The most important line of future research is that of the nature of the relation existing between emotion and learning, in the broadest sense of the term." P. 521.

[2] *Cf.* "Contributions on Learning" by Lashley and Hunter in "Foundations of Experimental Psychology," Chap. XIV and XV, respectively.

"Freud's concerted attack upon problems of the emotional life of human beings," may be, as Wheeler describes it, "the second important event of recent years"[1] in the field of psychology. Doubtless psychoanalysis has given us insights into human nature which scientific psychology working with narrow functions in artificial laboratory situations could not have perceived as rapidly, if at all. But to the foregoing quantitative questions psychoanalysis cannot give answers. In psychoanalytic literature, one finds some mysticism, some concern with ultimates, all explanatory principles and many revelations but no answers to these or any other quantitative problems. The characteristic method of psychoanalysis is classification, not measurement. In the psychoanalytic world things happen "universally" and "invariably."

Because they are summarized by Hunter and partly because they are not directly concerned with our present problem, experimental investigations of the law of effect, including the first two of Professor Thorndike's crucial experiments, will not be further commented on.

What a critical survey of the more direct experimental studies of the relations of feeling and memory reveals is described in some detail in a recent article.[2] For our present purposes it is enough to say that the need is not for either "enlightment" of or "disillusionment" with Freudian doctrines but the measurement of realities. A first essential in the carrying out of this suggestion is the restatement of the problem. The question for which the investigators have sought an answer has been: Does feeling—in the sense of pleasantness and unpleasantness—influence remembering or forgetting? This manner of stating the problem presupposes that feeling conditions memory. Or, as Weiss[3] expresses it, pleasantness and unpleasantness are treated as "something" which determines our action. As likely an explanation for associations found between these two processes is that much of what conditions learning also conditions our feeling reactions. To make this problem more susceptible to investigation and verification, the question would be better re-stated in terms which make neither of these assumptions. It would be better for the time being to waive interpretation and treat such data as the memory of pleasant experiences, the forgetting of pleasant experiences, the remembering of

[1] "The Science of Psychology." P. 23.

[2] Meltzer, H.: The Present Status of Experimental Studies on the Relationship of Feeling and Memory. *Psychological Review*, March, 1930.

[3] Wittenberg Symposium on "Feelings and Emotions." P. 177.

unpleasant experiences, the forgetting of unpleasant experiences as *unitary responses.*[1] In addition, to assure that realities are measured, the shortcomings of the previous studies should be avoided. Chief of these are: the assumption that humans experience equal amounts of pleasantness and unpleasantness; that predetermined stimuli can be used for evoking pleasant or unpleasant responses from all individuals examined; that the summation of discrete responses obtained by the use of the association method will yield results which are characteristic of total personalities; that reactions to such stimuli as pleasant and unpleasant odors in a laboratory are the same as feeling reactions in everyday life.

It is the awareness of these shortcomings which determined the direction of the present study. The purpose of this study, then, is to make a first attempt to determine the facts concerning individual differences in human forgetting of pleasant and unpleasant experiences, which are as life-like and personal, though not as astonishing as those reported in psychoanalytic literature. In this study, however, forgetting is limited in meaning to inability to recall and is not confused with the psychoanalytic use of the term "forget" in the sense of "Don't brood," "Don't worry about it," or "Dismiss from mind until you can view it in proper relation to rest of life experiences." Such instructions are not, as pointed out by Loveday,[2] equivalent to "do not endeavor to be unable to recall."

Procedure

One hundred thirty-two college students—seventy-seven men and fifty-five women—on a day following a Christmas vacation were asked to list and briefly to describe all their experiences of their vacation. They were then instructed to mark *P* before every experience they considered pleasant and *U* before every one they considered unpleasant. That done, they were further directed to rate every experience from the point of view of vividness[3] on a scale ranging from *A*, the most vivid to *F*, the least vivid. An *A*, *B*, *C*, *D*, *E*, *F* classification was used, largely because it corresponded with a distribution scheme known to the individuals because their marks were so

[1] In the Woodworth sense, not unanalyzable wholes but identifiable and measurable forms of realities.

[2] In symposium on The Rôle of Repression in Forgetting. *British Journal of Psychology*, Vol. VII, 1914–15, pp. 139–165.

[3] The rôle of vividness is to be reported in a later article.

reported. Six weeks later, without warning, the foregoing procedure was repeated. In this study we refer to the experiences given at first recall as experiencing and those given at the second recall as remembering. New experiences, that is those named on second time not named the first, were listed separately and eliminated from statistical study of experiences remembered and forgotten. In both instances the individuals were given all the time they wished to use. The nature of the methods used in treating these materials are revealed in the report of the findings.

Findings

1. *Differences in Experiencing Pleasant and Unpleasant.*—The total number of experiences reported by the one hundred thirty-two individuals on first recall was 2231. Of these 1393 or 62.43 per cent were marked pleasant, and 838 or 37.37 per cent were marked unpleasant. These same data treated more statistically yield the following results:

Table I.—Differences in Per Cent Pleasant and Unpleasant Experienced

	Per cent P experienced (P/T)	Per cent U experienced (U/T)
Mean	62.00	38.00
SD	11.00	10.00
N	132.00	132.00
Median	61.76	39.23
Q_1	54.60	30.22
Q_3	70.60	44.80
Q	8.00	7.29
Lowest	10.00	10.00
Highest	86.00	65.00
Ratio L and H	1:8.5	1:6.5

The foregoing results correspond to a surprising extent with results found by Flügel[1] who had nine observers keep account of their experiences for thirty days. The ratio he obtained was: 50.1 pleasant, 27.8 zero, and 22.2 unpleasant. If we break the zero grouping in two equal halves the results correspond very closely with ours. The two indicate a general tendency towards predominance of pleasant and our results also indicate large individual variations in such experiencing.

[1] A Quantitative Study of Feeling and Emotion in Everyday Life. *British Journal of Psychology*, 1925, Vol. XV, pp. 318–355.

Analyzed in terms of experiences per individual the results are: Mean number of pleasant experiences reported 11.06 (PE = ±2.81); mean number of unpleasant experiences reported 6.91 (PE = ±1.88).

This indicates that more pleasant than unpleasant are usually named. It follows, therefore, that differences in forgetting of pleasant and unpleasant experiences cannot be determined by tabulating and arithmetically treating the number of experiences; yet in many previous studies it has been tacitly assumed that we experience the same amount of pleasant and unpleasant experiences.

2. *Algedonic Differences.*—One of the means used by Flügel for determining the extent to which what he called pleasure and unpleasure occupy larger or smaller portions in human life he called by the name of algedonic difference. This is nothing more than the answer obtained by following the formula $P/T - U/T$ where P stands for pleasant experiences; U, for unpleasant; and T, for total. The average algedonic difference for his nine observers he reports as 27.9. Again our results correspond closely with his. This is indicated in the following table:

TABLE II.—ALGEDONIC DIFFERENCES

	First recall	Second recall
Mean	25.00	30.90
SD	19.50	27.00
Median	23.75	31.40
Q_1	10.50	13.40
Q_3	40.20	52.00
Q	14.85	19.30
Highest	+75.00	+90.00
Lowest	−35.00	−25.00

These results also indicate an increasing tendency with increasing time for P/T to be greater than U/T and for the variability also to increase.

3. *Differences in Forgetting Pleasant and Unpleasant Experiences.*—Of the 2231 experiences given on the first recall 1092 or 48.97 per cent were forgotten at the end of six weeks. Of the total pleasant experiences given the first time 593 or 42.06 per cent were forgotten and of the total unpleasant, 499 experiences or 59.53 per cent were forgotten. A more statistical analysis of these data presented in the next table indicates the extent and significance of the difference between the forgetting of pleasant and unpleasant experiences:

TABLE III.—DIFFERENCES IN PLEASANT AND UNPLEASANT FORGOTTEN

Mean number	PE	Mean per cent	SD dist.	PE_m	Highest	Lowest
P_f, 5.45	1.76	46.97	18.45	±1.13	0	100
U_f, 4.28	1.64	60.25	25.38	±1.56	0	100

Difference in favor of U = 13.28
PE difference = 1.93
D/PE diff. = 6.88

The foregoing results indicate that more pleasant than unpleasant experiences are forgotten but that in per cent forgotten there is a statistically significant difference of 13.28 in favor of the unpleasant.

The large individual variations in both the forgetting of pleasant as well as unpleasant experiences is indicated in the following frequency distributions:

TABLE IV.—FREQUENCY DISTRIBUTION OF PER CENT PLEASANT AND UNPLEASANT FORGOTTEN

Per cent experiences	Frequency of pleasant	Frequency of unpleasant
100	2	13
95	..	..
90	..	1
85	..	4
80	1	11
75	5	11
70	6	6
65	4	17
60	12	9
55	12	11
50	18	14
45	10	..
40	18	4
35	11	4
30	12	6
25	8	9
20	4	3
15	4	3
10	..	1
5	..	..
0	5	5
Total	132	132
	Mean = 46.97 SD = 18.45	Mean = 60.25 SD = 25.37

The psycho-social factors which have played a rôle in the foregoing results are many in number and complex in nature. The present study was not set up to determine these. Some of the factors, however, are apparent in the present data. One is the fact that, as Ballard found, in his study of "Obliviscence and Reminiscence,"[1] "We not only tend to forget what we have remembered but also tend to remember what we have once forgotten." That is, "Obliviscence is at least partly conteracted by an opposing tendency toward reminiscence." A rough measure of the extent of this counteracting tendency present is the number of new experiences, those reported the second time but not given the first. The mean of new experiences reported are as follows: $1.42 \pm .95$ pleasant experiences; $1.12 \pm .75$ unpleasant experiences; and 2.12 ± 1.43 total experiences. A likely explanation for this tendency is that given by Crossland[2] after direct experimentation with such phenomena. It is to the effect that attempts to overcome errors of omission result in errors of commission or elaboration.

In the forgetting of individuals who not only forgive everything but forget everything, whether pleasant or unpleasant, stupidity doubtless plays a rôle.

The rate of forgetting of personal experiences is less than the rates for the forgetting of nonsense syllables, poetry and other materials reported in literature.[3] The means for pleasant and unpleasant experiences forgotten have been previously reported. For total personal experiences forgotten in six weeks the mean is 51.48 and the SD is 25.38. Pear's[4] explanation for this difference is in the greater "affective relevance" of personal experiences. Herrick's[5] explanation in terms of relative meaningfulness associated with greater intensity is in essence similar.

4. *Differences in Indices of Feeling Tones.*—From the first to the second recall, *i.e.*, with the passing of six weeks time, what change takes place in the ratio of pleasant to total remembered? Of unpleasant? How large are the individual variations in both instances? On the

---

[1] In British Journal Psychological Monographs Supplement Vol. I, 1913.

[2] A Qualitative Analysis of the Process of Forgetting. *Psychological Monographs*, 1921, Vol. XXIX, No. 130.

[3] *Cf.* Hunter's contribution to "Foundations of Experimental Psychology." P. 606.

[4] "Remembering and Forgetting." P. 138.

[5] "The Thinking Machine." P. 224.

average, which tends to increase more with time, the tendency of the pleasant tone or unpleasant tone? To answer these questions we have calculated a pleasant tone index, an unpleasant tone index, a pleasant-unpleasant potency index and a total feeling tone index for each individual. The formulas used in these determinations were:

Pleasant tone index = $P_2/T_2 - P_1/T_1$.
+ indicates increase with time in the per cent of pleasant memories.
0 indicates per cent same as first recall.
− indicates decrease with time in per cent of pleasant memories.
Unpleasant tone index = $U_2/T_2 - U_1/T_2$ (similar to P tone).
$P - U$ potency index = P tone index − U tone index.
(+ indicates U more forgotton; 0, equal potency; − P more forgotten.)
Total feeling tone = pleasant tone index + unpleasant tone index.

The results obtained by the use of these indices are shown in the following table:

TABLE V.—DIFFERENCES IN FEELING TONE INDICES

	Pleasant tone	Unpleasant tone	$P - U$ potency	Total feeling
Mean	+4.79	−3.68	+9.40	20.08
PE	±8.73	±8.81	±17.43	±12.07
Highest	+50.00	+20.00	100.00	105.00
Lowest	−25.00	−55.00	−50.00	0.00

SD = 25.8
Median = 6.6
Q_1 = −7.5
Q_3 = 24.2

75 per cent P tone indices exceed mean of U tone

In all the foregoing indices large variations are the rule. The extent of the differences on pleasant tone index and unpleasant index is indicated in Fig. 1.

The extent of variation on total feeling tone is indicated in Fig. 2.

The $P - U$ potency index comes nearest of all measures used in this study to being an indicator of the inhibition of the unpleasant. The distribution on it is given in Fig. 3.

In the foregoing distribution there is nothing that indicates that humans invariably tend to inhibit the unpleasant. What it does show is: (1) Large individual differences in forgetting the pleasant as well as the unpleasant; (2) the existence of many individuals who tend to forget the pleasant more than unpleasant; and (3) that in the

group studied there are more individuals who forget the unpleasant more and to a larger extent than the pleasant.

If we divide the $P - U$ potency distribution into three groups, we get in the first two (1) individuals who tend to forget the unpleasant more than the pleasant (who might be said to be optimistically-minded) and (2) individuals who tend to forget the pleasant more than the unpleasant (who might be called pessimists). In the third group are

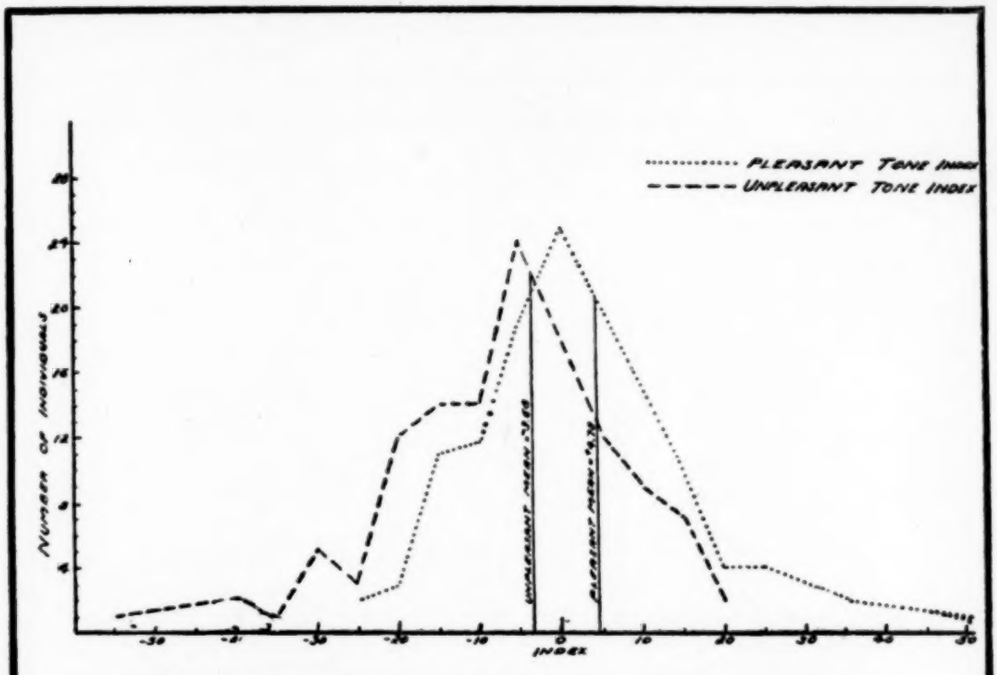

FIG. 1: SHOWING DIFFERENCES IN DISTRIBUTIONS ON INDICES OF PLEASANT TONE AND UNPLEASANT TONE

the indifferentists, who tend to forget pleasant and unpleasant in equal proportions. In this last group are only those individuals whose $P - U$ potency index = 0. The distribution of the total group is then as follows:

Class	Per Cent
Optimists	56.49
Pessimists	35.87
Indifferentists	7.64
Total	100.00

Summary of Findings

1. A total of 2231 experiences were named on first recall. Of these 1393 or 62.43 per cent were marked pleasant and 838 or 37.5 per cent were marked unpleasant.

2. The extent of individuals differences in per cent pleasant and unpleasant experienced is reported statistically in the body of the paper.

3. More pleasant than unpleasant experiences are named on the average. The mean number of pleasant experiences is 11.06 ± 2.81. The mean number of unpleasant experiences reported is 6.91 ± 1.88. Yet in many previous studies it has been tacitly assumed that one experiences the pleasant and unpleasant in equal amounts.

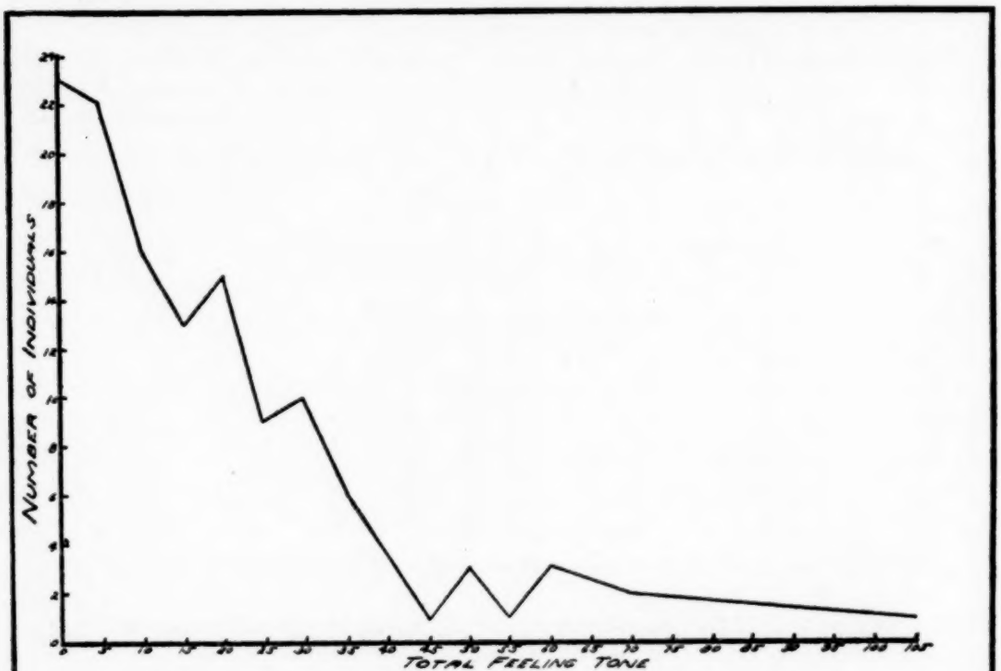

FIG. 2: SHOWING DISTRIBUTION ON INDEX MEASURING ROLE OF TOTAL FEELING TONE IN FORGETTING

4. The algedonic differences, the tendency for P/T to be larger than U/T, agree with Flügel's findings obtained by an entirely different procedure.

5. After a lapse of six weeks, 48.97 per cent of total experiences are forgotten; 59.73 per cent of the unpleasant; and 42.60 per cent of the pleasant.

6. The mean number of new experiences, those given in second recall though not given in first, are as follows: 1.42 ± .95 pleasant; 1.12 ± .75 unpleasant; and 2.12 ± 1.43 total experiences. The most likely explanation for the appearance of these experiences is that given by Ballard in his study of reminiscence; namely, that we

tend not only to forget what we once remembered, but that we also tend to remember what we had forgotten.

7. There is a statistically significant difference of 13.28 per cent in favor of forgetting of unpleasant. (*PE* differences = 1.93 and $D/PE = 6.88$.)

8. The rate of forgetting for personal experiences is less rapid than for other materials reported in literature.

9. A study in differences in pleasant tone indices ($P_2/T_2 - P_1/T_1$), unpleasant tone indices ($U_2/T_2 - U_1/T_1$), $P - U$ potency indices (P tone $- U$ tone) and total feeling tone, show wide differences in individual responses but that in the group studied there are more individuals who forget the unpleasant and do so to a larger extent than individuals who forget the pleasant more.

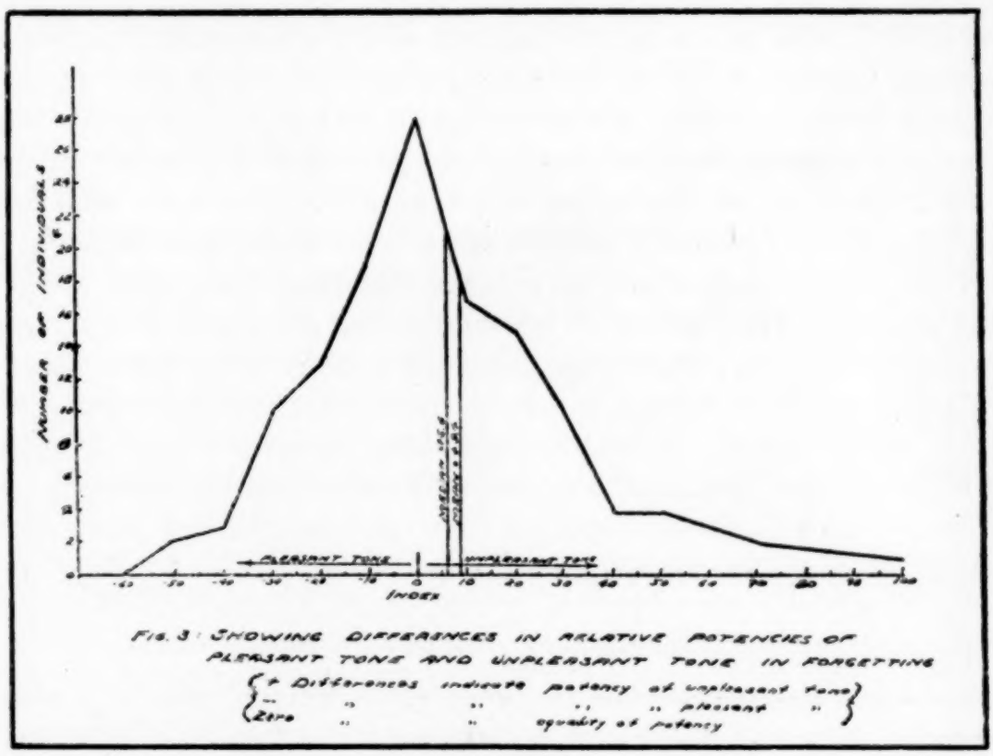

FIG. 3. SHOWING DIFFERENCES IN RELATIVE POTENCIES OF PLEASANT TONE AND UNPLEASANT TONE IN FORGETTING

{+ Differences indicate potency of unpleasant tone
− " " " " pleasant "
Zero " " equality of potency}

10. If we call individuals who tend to forget unpleasant experiences more than pleasant, optimists; those who tend to forget pleasant experiences more than unpleasant, pessimists: and those who forget both in equal proportions, indifferentists; our group classifies itself as follows: 56.49 per cent optimists, 35.87 per cent pessimists, and 7.64 per cent indifferentists.

THE INFLUENCE OF ATTITUDE UPON TRANSFER

MATTIE F. DORSEY

University of Colorado

AND

L. THOMAS HOPKINS

The Lincoln School of Teachers College, Columbia University

The purpose of this paper is to report the results of an experiment to determine the variations in the amount of transfer depending upon readiness to act as built up by suggestion given at the time of performance. The experiment consisted of three separate divisions, or cases, in each of which the equivalent group technique was employed. To each group of a case was furnished an identical situation to which it was possible to transfer elements of previous experience. To the experimental group was given the suggestion that such experience might be used in meeting the new situations, whereas no such suggestion was given to the control group. Identical tests were given to both groups. The customary training period used so widely in transfer experiments was unnecessary since the members of each group in each case had been taught the same basic materials and methods by the same individual. It was assumed that transfer would take place in both groups and that whatever difference should occur in performance between the two groups of a case would be due primarily to the influence of the experimental factor.

Materials and Methods

Case I was composed of ten boys and eighteen girls in a college freshman class who had failed in their work the first quarter and who were given special instruction in how to study for the second quarter. The factor to be transferred was *method of study* to the twenty-four minute reading of test materials composed of ten paragraphs in general psychology. Measurement was based upon a ten minute uncontrolled reproduction of the ideas presented and a true-false examination covering the same materials. Equating was made by alternate advantage using scores on a college reading test. To the experimental group was given the suggestion: "Use the method of study which you have learned in studying these paragraphs."

Case II was made up of twenty freshman girls who were taking first year college Latin, based upon four years of study of this subject

in secondary school. The factor to be transferred was *knowledge of Latin* as applied to a one-out-of-five multiple-choice test containing one hundred Latin derivatives, fifty from words previously learned and fifty from new words. The groups of ten students each were equated on the basis of composite scores from a group intelligence test, grades in Latin for two quarters and scores on the Stevenson-Coxe Latin Test. To the experimental group was given the suggestion: "Use your knowledge of Latin in determining the meanings of the words in the test."

Case III included one hundred freshmen engineering men who had just completed the study of descriptive geometry under the same teacher. The factor to be transferred was *skill in manipulation of special elements* to the new situation of a speed test consisting of (1) ten problems in short division to be solved mentally, (2) the straight line alphabet, (3) a geometry test of counting the number of lines necessary to construct certain objects described in the test, (4) the cube test, (5) writing as many words possible in five minutes from the letters of the word "material" which had been given orally. The groups were equated upon a number of individual and multiple bases which will be subsequently described. At the beginning of the test the experimental group was given the suggestion: "Use your knowledge of descriptive geometry in answering the test questions."

All the students in each group were taught by the same teacher, making for uniformity in content and method. All of the tests were constructed, given and scored by two persons who were familiar with test techniques and experimental research. In each case the two groups took the test at the same time, thus preventing carryover from one group to another.

Group Data

Case I.—The experimental factor in Case I was method of study. The experimental subjects were given two tests based upon a twenty-four minute study of ten paragraphs of material in general psychology. A ten minute uncontrolled reproduction test was scored by counting the number of words written. The true-false test was scored in two ways: (1) One point of score was given for each correct answer; (2) correction was made for guessing by subtracting the wrong responses from the right.

Table I presents the total and mean scores, standard deviation, per cent of overlapping and experimental coefficient for the control

and experimental groups in Case I. These data seem to warrant the following inferences:

1. The total and mean scores of the experimental group are superior to those of the control group on both tests. For the true-false test the corrected score yielded greater differences than the uncorrected score.

2. The experimental group has a greater standard deviation than the control group indicating that the experimental factor caused greater variability within the group.

3. The per cent of cases of one group overlapping the mean of the other is in favor of the experimental group with a range of 14.3 to 42.9 per cent.

TABLE I.—SUMMARY OF FINDINGS FOR CASE I

	Number words written	True-false scores, rights	True-false scores, rights-wrongs
Total Scores			
Control	1295	261	178
Experimental	1496	284	215
Difference	201	23	37
Mean Scores			
Control	185	37.28	25.42
Experimental	213.7	40.57	30.71
Difference	28.7	3.29	5.29
Gain per cent in favor of experimental group	15.5	8.8	20.8
Standard Deviations			
Control	30.57	4.19	8.44
Experimental	47.09	5.01	9.58
Difference	16.42	.82	1.14
Per Cent of One Group Overlapping Mean of Other			
Control	14.2	28.5	57.1
Experimental	57.1	71.4	71.4
Difference	42.9	42.9	14.3
Reliability of Difference between Means			
Experimental, coefficient	2	1.99	5.17
Chances in 100, difference is above zero	91	91	100

4. The experimental coefficient found by dividing the difference between the two means by the probable error of the difference indicates that the chances are from ninety one to one hundred in one hundred that there is a true difference in the performance of the two groups.

Case II.—Table II presents the average per cent scores, standard deviation, per cent of overlapping and the experimental coefficient for the control and experimental groups in Case II. This information is presented under three headings—total words, old words, and new words. The test contained fifty English derivatives from known Latin words and fifty English derivatives from words which had not been studied in class. This division was made to determine the transfer effect of the readiness suggestion with old and new materials. From the data presented in this table are drawn the following conclusions:

1. The experimental group is superior to the control group in the average total number of words defined correctly.

TABLE II.—SUMMARY OF FINDINGS FOR CASE II

	Total words	Old words	New words
Mean Scores Expressed in Percentages			
Control group	63	68.8	40.5
Experimental	69.1	75.4	54.5
Difference	6.1	6.6	13.9
Gain per cent in favor of experimental group	9.6	9.5	34.4
Standard Deviations of Percentages Correct			
Control group	4.69	5.	11.6
Experimental group	9.50	12.2	12.1
Difference	4.81	7.2	.5
Per Cent of One Group Overlapping Mean of Other			
Control group	10	10	10
Experimental group	70	70	80
Difference	60	60	70
Reliability of Difference between Means			
Experimental coefficient	2.71	2.35	3.89
Chances in 100 difference is above zero	97	94.5	99.6

2. The experimental group made greater gains over the control group with the new words than with the old words, the difference being 3.4 per cent in the former and 9.5 per cent in the latter.

3. The standard deviation of the experimental group is higher in each case than for the control group although it is smallest with new words. This seems to indicate that the readiness suggestion produced greater individual differences with old words than with new words. In the latter case, it tended to keep the two groups uniform in spread although one was far superior to the other in average performance.

4. The experimental group shows decided superiority over the control group in the per cent of cases overlapping the mean of the other group. The difference is from sixty to seventy per cent.

5. The experimental coefficient is large enough to indicate from 94.5 to 99.6 chances in one hundred that the difference in the mean is significant and that the experimental group is actually superior to the control group.

Case III.—The experimental factor in Case III was skill in the manipulation of spacial elements. The experimental subjects had just completed the study of descriptive geometry under the same teacher. The transfer situation was a speed test consisting of five separate parts described above. The results are presented in Table III under six different methods of grouping made possible by the larger number of subjects than in the preceding cases and by the absence of equating until after the testing. These groups are as follows:

1. The entire control group of forty-seven compared with the entire experimental group of fifty-two members.

2. Two groups of thirty-seven members each, equated on the basis of grades in descriptive geometry.

3. Two groups of twenty-eight members each, equivalent in ability in descriptive geometry and in a college classification test.

4. Two groups of freshmen equated on the basis of composite scores obtained by combining scores on the Iowa Placement Tests with final grades in descriptive geometry. Three methods of weighting were used for making these scores and grades comparable hence there are three sub-groupings under 4.

(*a*) Composite scores obtained by weighting according to percentile ranks.

(*b*) Composite scores obtained by weighting according to variability, using the standard deviation.

(*c*) Composite scores obtained by weighting according to standard scores, *i.e.*, raw score minus the mean divided by the standard deviation.

The groupings are arranged according to degree of refinement in technique of equating beginning with Column 2 in which only one basis of equating is used and increasing to Column 4*c* in which grades in geometry, classification in college and scores in three Iowa Place-

ment Tests are used in order to obtain a composite score, and in which these are weighted according to a more reliable method of making scores comparable.

The data presented in Table III serve as a basis for drawing the following conclusions:

1. The total and mean scores of the experimental group are higher than those of the control group with every method of grouping.
2. The standard deviation is less for the experimental than for the control group in five out of the six classifications indicating that the experimental factor tended to make the experimental group less variable than the control group.

TABLE III.—SUMMARY OF FINDINGS FOR CASE III

	Groupings					
	1	2	3	4*a*	4*b*	4*c*
Total Scores						
Control group	1366	994	768	550	596	527
Experimental group	1619	1146	882	649	702	654
Difference	253	152	114	99	106	127
Mean Scores						
Control group	29.95	27.67	28.42	28.60	28.18	27.40
Experimental group	31.61	31.67	32.14	33.20	32.72	33.60
Difference	1.66	4.00	3.72	4.60	4.54	6.20
Gain per cent in favor of experimental group	4.2	14.45	13.09	16.04	16.11	22.62
Standard Deviations						
Control group	1.98	1.66	1.88	2.12	2.07	1.90
Experimental group	1.60	1.80	1.67	1.73	1.61	1.74
Difference	−.38	.14	−.21	−.39	−.46	−.16
Per Cent of One Group Overlapping Mean of Other						
Control group	39	28	31	29	28	17
Experimental group	56	69	71	73	76	81
Difference	17	41	40	44	48	64
Reliability of Difference between Means						
Experimental coefficient	7.2	15.38	12	11.61	12.27	16.31
Chances in 100, difference is above zero	100	100	100	100	100	100

3. The per cent of overlapping is greater for the experimental than for the control group. This difference increased with the reliability of the method of equating groups.

4. In every instance the experimental coefficient is high enough to indicate with certainty that the difference between the two groups is real and greater than zero.

Individual Performance

The measures of variation used for comparison indicate great individual difference within the groups. Examination of individual scores reveals the following facts:

1. In no comparison in any of the three cases did every member of the experimental group exceed the member of the control group with whom he was paired.

2. There are no instances in which there was a greater per cent of the experimental group members who failed to exceed their partners in performance, than of the control group.

3. In one comparison in Case I and in four of Case III there were pairs in which neither member exceeded the other.

4. In Case I there was no pair in which the control member made a higher score in the true-false test than his partner.

5. The range of differences in favor of the experimental member of a pair extended from forty-one per cent to one hundred seventy-five per cent, for the whole experiment.

6. In Case I the greatest variation was shown in corrected true-false scores where the gain in favor of the experimental member ranged from 12.5 to ninety per cent. In Case II the range of excess was −11.6 to 83.1 per cent for experimental members. Case III, Grouping 3 showed a variation in gain per cent in favor of the experimental members of forty-one to one hundred seventy-five.

7. In Case II the differences between performance of the members of a pair varied with different comparisions, such as between old words and new words correct.

From these facts, it seems safe to conclude that (1) attitude causes individuals to vary in amount of transfer not only in the group to which they belong, but also in comparison with individuals with whom they have been paired and (2) attitude affects transfer differently for the same individuals under varying conditions.

Summary and Conclusions

1. This study reports group and individual data from an equivalent group experiment with college freshmen to determine the amount of transfer depending upon readiness to act as built up by suggestion given at the time of performance. The experiment consisted of three separate divisions or cases. Case I involved the transfer of *method*

of study to the study of a particular kind of material. Case II involved a transfer of *knowledge* from a subject to an allied field containing new and old materials. Case III involved a transfer of *skill* in the manipulation of ideas from the situation in which it was learned to a new one outside of the field.

2. Total scores, mean scores, per cent of one group overlapping the mean of the other and the experimental coefficient in all three cases indicated marked superiority in the performance of the experimental over the control groups.

3. Standard deviations indicated that with Cases I and II the experimental factor tended to produce greater individual differences and spread the performance of the experimental group over the control group. Exactly the opposite condition was found in Case III and may be due to a larger number of cases and different types of material or to more reliable methods of equating.

4. Since the experimental and control groups in each case were equated on the basis of intelligence, the superiority of the experimental group over the control group cannot be due to ability, but must result from some other influence such as readiness to act.

5. Since the two groups in each case were given the same basic training in the factor to be transferred, the difference in amount between the two groups must be due to the presence of some factor other than previous training.

6. The mere presence of a possible transfer situation would produce some, but not the maximum amount of transfer.

7. The amount of transfer both with and without the readiness factor is an individual matter. Paired subjects having the same ability, the same previous training, the same suggestion or lack of suggestion, and the same test differ in the amount of transfer.

HABIT AND ATTITUDE

HAROLD SAXE TUTTLE

University of Oregon

Recent studies of behavior demand a redefinition of "habit." The use of the word in the past has been ambiguous. Considerable confusion has resulted.

As commonly defined, habit means the tendency under given circumstances to behave in fixed ways. By one group of writers it is described in terms of economy of effort—the line of least resistance. This definition emphasizes the now familiar fact that the physical basis of habit is to be found in the organization of neurones leading to specific muscular responses. When a given stimulus appears a second time it tends to follow precisely the same neural arc arousing identically the same muscular response. This neural arc represents the line of least resistance.

Unless all our knowledge of physiology is a myth this is basically a true description of habit. Motor responses are the same under a given environment because the stimulus finds least resistance in following the established connections through the synapses involved, although other potential alternatives exist but offer greater resistance. This is habit.

The significant aspect of habit, defined in terms of synaptic connections, is its economy of energy. Large areas of life may be reduced to habit, releasing attention for the critical situations which can not thus be cared for. Dressing, eating, walking, driving, writing may all be accomplished with consciousness directed the while to the new, the unusual, the challenging. The mechanic performs complicated tasks with apparently no attention to his work. One may enter into a spirited conversation while he eats or drives; he may compose significant statements as he writes; he may note the beautiful sunset as he walks. The enormous economy of energy made possible by habituated conduct is due to its basic dependence upon unresisting nerve pathways. Habit is life's greatest thrift examplar. All this is possible because recurring stimuli follow synaptic lines of least resistance

If this were the last word to be said about habit the whole matter would be reasonably simple. Every parent and teacher who would take the trouble to read a simple description of the nerve system

could understand not only the practical workings of habit but the scientific explanation of the process.

But this is not the last word. There is another and quite different element in all the definitions from James to the present moment. The contrast may perhaps be seen by asking the question: Is habit a line of least resistance, a least-effort escape from a stimulus, or a dynamic tendency, a slave driver that will not be disobeyed?

The second group of writers (or rather of definitions, for the authorities are in most cases the same for both positions) emphasizes the dynamic element in habit. Habit is motive. It is the preferred way of responding to a given situation. To respond otherwise is objectionable. The familiar way of behaving carries with it an imperative that brands alternatives as not-to-be-thought-of.

The boy who has developed the habit of thrift is unwilling to spend until he has put some of his earnings in his bank. The truthful child would recoil at the thought of telling a falsehood. In old age the mother-in-law is distracted by the methods of cooking and housekeeping practiced by the younger generation. A girl is not credited with the habit of neatness unless she is uncomfortable in the presence of disorder. And even more evident is the slavery to *bad* habits—which need not be recited here in order to carry conviction.

The "folkways" dominate successive generations because they are cultivated in children too young to analyze their worth, and then practiced without reflective thought because they are habitual and have become imperative. Traditionalism would not be so enslaving were it not for the dynamic nature of habit.

Society disagrees as to whether the duty devolves upon home, church, or school, but all agree that somewhere there rests the duty to train children in habits of courtesy, truthfulness, obedience, punctuality, orderliness, honesty, and many like virtues. In this assumption there is implicit a definition of habit as something dynamic and imperious, not merely a line of least resistance.

"The hell to be endured hereafter, of which theology tells, is no worse," wrote William James in 1890, "than the hell we make for ourselves in this world by habitually fashioning our characters in the wrong way. Every smallest stroke of virtue or of vice leaves its never-so-little scar." And twenty years later Bagley wrote: "Prejudices and attitudes may grow out of specific habits, as when the habits of Sunday observance, established in early childhood, become more or less explicitly formulated as ideals and gradually come to

express themselves as prejudices which make the lack of observance a matter of annoyance."

In 1927 Dashiell wrote in his discussion of "habit": "Interruption or thwarting of a routine manner of conduct frequently acts upon the individual as a positive irritant." And again: "No matter how healthful, cheap, convenient and attractive a new article may be, prospective purchasers will show a pronounced tendency to continue to buy the old familiar ones."

Now, with these two elements in the description of habit set side by side, it must be apparent that they do not refer to precisely the same types of conduct. The line of least resistance leads through synapses to muscles. This defines the form of motor response. It is most clearly seen in skills. Each repetition of the appropriate situation calls out with precision the appropriate response. The finger touches the key or moves the pencil or the chisel in the prescribed way more quickly than consciousness can change its focus. Here is a perfect exhibition of habit as a response in the line of least resistance.

The process may be extended into a complicated series of motor responses, yet remain the same psychologically. A habit need not be simple in order to justify explanation as a definite response to a familiar stimulus operating through familiar synaptic connections. But to claim such a definition, it must involve a *fixed form of motor response.* The $S \rightarrow R$ may be extended by reciprocal innervation into a long sequence of $S \rightarrow R$ provided that the R's are definite and uniform in each repetition.

This is not rigidly true of the habits described in the second series of definitions. And if not rigidly true, the explanation of synaptic pathways loses its force altogether. What is the nature of the habits which have been defined for the past forty years as dynamic, as controls, as motives, as drives? With James it is drinking; with Dashiell it is buying advertised goods. And the list might be extended monotonously. Eating at noon, turning to the right, telling the truth, keeping desks neat, showing courtesy, singing in school assembly, driving carefully—these are but a few of the habits which, with their opposites, come in for a place in that training which will make them either second nature or anathema.

Now, not one act in the list from James to Dashiell is a precise motor response. Not one of them need, so far as the definition goes, occur twice in exactly the same way muscularly. Not one can be

strictly accounted for (in which case the word *strictly* may be omitted for scientifically minded thinkers) in terms of fixed neural pathways.

In every instance the *outcome* is the determining factor. In each case the dynamic element is associated with the result. *Previous experience is always implied, and that experience always involves an element of satisfaction.*

The feelings have functioned in association with the outcome. A sense of worth has been established as an integral part of the response. Food tasted good; traffic was successfully negotiated; pride has been shown at the courage which inspired the truth; the neatness of the desk was praised; the music in assembly gave everyone a sense of loyalty to the school. And these results have occurred repeatedly. Interest in the outcome has developed all but unconsciously until the noon bell or whistle makes eating highly desirable; not to turn to the right would be an inexcusable discourtesy; an untruth is simply not to be considered; the desk must be neat no matter how great the hurry. These "interests," when fairly definite, we have come to call habits. But when subjected to precise analysis it is clear that they do not conform to the first type of definition, they are not motor responses dependent on established neural connections of least resistance. Such behavior is accounted for not in terms of repetition but in terms of satisfaction.

Indeed, such conduct involves no regularity of form or order of motor responses. Perhaps they are almost the same each time; often they are essentially uniform; but if one will observe precisely it will be evident that the muscular movements involved vary in kind and in order. The determining element in the behavior is not the "pathway" in the brain but the goal in conduct.

If a pioneer farmer were to haul a crate of dynamite to his stump-filled field and blow some of the stumps, then burn others and pull still others, one would hardly say he had the "habit" of clearing stumps. When some element of worth, of interest, of desire, related to the service rendered by any particular response, determines the form of that response, it is essentially an *attitude* and not a habit.

Virtually all the ordinary habits of life are colored by some affective state. Dashiell's shoppers could hardly be accused of stoic indifference. Bagley's Sunday observer is frankly credited with prejudice. James' Rip Van Winkle and all his spiritual kin make no secret about the feelings they experience toward liquor. The tobacco "habit," the drug "habit," the gambling "habit" are cases of affective attitudes

toward "values" in suggested responses, not drives of a stimulus through a fixed neural arc. No more precise or uniform can be the conduct exemplifying "habits" of courtesy, truthfulness, obedience or chivalry.

Protests that "general habits" are not unities, even if true, do not clarify the situation greatly. No part of the total behavior which constitutes the "general habit" is any more strictly a habit than the total. One does not behave alike in any physiological sense when he returns a dime to the owner who has unconsciously dropped it and when he avoids taking the quarter in plain view in an open coat pocket in the coat room. Motor repetition cannot explain his haste to be punctual Thursday when he had a late start, after having walked deliberately on Wednesday when he had ample time to reach school. No two acts of virtue, however specific, are based on a fixed neural arc.

There are certain rigidly exact motor skills. Ordinarily these are thought of in connection with arts, esthetic or practical. Some such skills are involved in everyday behavior, as the pronouncing of words, spelling, penmanship, manipulations in driving a car and the like. But these are not complete responses to social situations. They do not constitute conduct in the sense implied by the ordinary use of the word "habit." And apart from these there are no strictly controlled motor responses in such everyday behavior as we commonly class under habits. We are dealing, after all, with attitudes, interests, traits, and ideals, rather than with "pure" habits.

Nor would it be accurate to distinguish the two forms of responses as dependent one on the "law of use" and the other on the "law of effect." Mere repetition (use) would lead, of course, only to the most exact following of neural arcs in a fixed order. But, strictly speaking, the association of satisfaction with a given response exemplifies the "law of effect" only when it strengthens the tendency of the given stimulus to follow *through the same arc to the same specific motor response.* The rat follows a given pathway in the maze leading to food (satisfaction). And when that series of motor reactions has become habitual through satisfactory outcomes he continues the same series even when a glass top admits light plainly showing a partition which has been placed in the middle of the long alley. Bumping into the visible partition indicates a series of fixed motor responses established by previous satisfactions. The kitten jumps through the hoop as the only route by which to get her milk until finally she jumps through the hoop even when there is ample room beneath it. Here we see the

law of effect building habits. Satisfaction strengthens the bond between stimulus and response. A fixed arc is utilized exactly as in the case of habits based upon use.

But the child who is praised for his thoughtfulness when he picks up crumpled papers from the floor is not utilizing a fixed neural bond between stimulus and response when he picks up unsightly objects from the school ground, or returns the pencil which the teacher dropped. He is adapting a method to a result which is "thoughtful" because the thoughtfulness brought approbation.

When truth-telling is praised, no matter how specific may be the situation, it is impossible that approval shall lead to another identical motor response in an exactly similar situation. It cannot fix a series of synaptic connections favoring a certain outcome. It can only set up a favorable affective state in connection with the *social value of truthfulness;* the means of attaining truthfulness in each subsequent case must involve some slight degree of reflective thinking.

Nor is it less true with regard to obedience. The form is constantly varying. The means of accomplishing the result, obedience, are continuously different. Some comparison must be made in image patterns which will anticipate the success of the contemplated response in realizing the value with which satisfaction has been associated. To some degree affective states and reflective thinking are involved in all these habits which have social significance. So far as moral conduct is concerned, there are no genuine habits. The outcome always to some degree determines the method chosen to secure it.

The study of mathematics cannot produce habits of accuracy for the simple reason that accuracy is an abstract description of a social value. The habit of writing 48 after the symbols 6×8 may be established. But when 9×7 is presented a different series of synaptic connections is involved. And so with each specific mathematical response. There is no neural arc which leads from the stimulus "mathematics" to the motor response, "accuracy." Accuracy is an ideal. It involves the feelings. It wins its influence in behavior by the satisfaction which comes from social approval of "correct" computations, the economic advantages of reliability, or the inner joy of victory over a hard problem. In no psychological sense can it be called a habit. The means by which it is to be attained call for choice, for comparison of data, for reasoning.

Two different words, therefore, are needed to describe responses. If habit is to be defined in terms of $S \rightarrow R$ then social behavior is not

a sum of habits. If profanity is a habit, then the $S \rightarrow R$ theory does not explain habit; for the man of forty who has indulged in profanity for thirty years and has never touched a hammer, will promptly swear when he strikes his thumb, although that stimulus has never before been associated with certain contractions of vocal chords and lungs. When we are seeking to guide responses by regulating stimuli, we need a word which clearly, exactly, and adequately indicates the type of neural function involved. When we seek to guide responses by associating satisfactions with the outcomes, where there are several alternate methods of securing the outcomes, we need a word which describes this function in contrast with former.

This distinction throws light upon the dual nature of what has been classified as habit. All that has been emphasized as to the dynamic nature of habit by James and all his successors is eminently true,—of an aspect of human conduct, which we shall continue to speak of as habit. All that Thorndike and his followers have said about the dependence of habit on synaptic paths is also true. The $S \rightarrow R$ formula is true and significant. The law of use and the law of effect describe tendencies to repeat a motor response set up by a given stimulating situation when repeated frequently or accompanied by satisfaction. The significant fact is that the types of behavior described in these two ways are not the same. This explains a difference between the training of children and animals. Belasco's cat could be trained to enter the stage, cross to the fireplace, stretch and yawn, then walk off stage through the other door. But cats have not been trained to such habits of courtesy or obedience or accuracy as a child can form. This explains also the hardship and annoyance involved in forcing an old person to change his ways of life. And by the same tokens it explains the dead weight of tradition upon all advances in civilization. To avoid disturbing the soil where evil spirits are wont to dwell does not involve a specific sensori-motor arc. But it does involve an affective attitude in connection with the outcome. To defend the sacredness of private property does not involve any fixed stimulus-response bond. But it does involve a feeling of worth concerning a definite adjustment.

Most significant of all is the light which this distinction throws upon the problem of conduct control in such fields as citizenship training and character education.

A very wholesome revolution has taken place within the last quarter century; very largely, indeed, within the past decade. Moral training

has been taken out of the realm of sentiment and subjected to the careful scrutiny of the psychologist. Such a step was necessary for genuine progress. It represents the first really promising movement toward meeting the perennial youth problem that has been offered since the dawn of civilization.

The essential contribution which has thus far been made is the rejection of vague, general ideals as objectives and the substitution of specific goals for immediate training. In one of his lectures popularizing psychology Overstreet pointed out that influencing behavior seemed like a very mysterious process, but that cultivating specific habits offered no such baffling difficulties; yet that the first is but a series of steps of the second sort. G. B. Watson in a stimulating article has insisted that character can not be trained by the cultivation of traits, since these are not unities; but that habits alone can be cultivated and in a sense controlled by others.

The specific habits referred to by these and many other writers are indeed much more sharply defined than the vague ideals of truthfulness and chivalry and all the rest. But are they habits of the $S \rightarrow R$ type? Are the motor responses in specific acts of respect for property ever twice alike? Is reaching out the arm for an apple the socially wrong act? Is suppressing the impulse to reach for the apple the socially good act? Clearly the "habit" to be formed—even the act to be performed—involves adjustment to a wide and complex range of elements other than sensory stimuli. Values are involved. Whether the "grocer wouldn't like it," or the "police would take you to jail," or merely "you have to pay before you take things in stores," the focus of attention is not upon the form of the motor responses reaching for the apple or suppressing the impulse, but upon the social value of the outcome. An element of knowledge of how to secure the desirable quality of the outcome is necessary.

So long as the babe depends on "yes" and "no" for every new situation and avoids only specific things or specific motor responses in a given situation, his conduct has only potential social value. His habits are quite inadequate for the needs of society. Not until he has organized sufficient experiences to understand what outcomes are to be secured, what values conserved, what purpose accomplished by certain responses, can he express his "habits" in varying situations. Only then can he understand explanations as to why he should do or avoid this act or that. His significant habits, then, are not the fixed responses of innocent babyhood, but the fidelity of his conformity

to recognized *values*. Most of his acts, to be sure, are so similar as not to appear variable. Yet the assumption of sensori-motor bases of behavior will not hold when rigidly scrutinized.

If it is contended that the value of the act, the previous approval of the response, as conditioning elements constitute a part of the total situation which stimulates the response, and therefore leave the act technically within the more precise definition of habit, two replies must be made:

1. The affective state is not a sense stimulus and can not have motivation value for a motor response unless the neural arc is definitely the same as in former cases; that is, unless the act is a habit in the narrower, more precise sense. And the opposite is true by definition; it is the behavior associated with values in unfamiliar outcomes that constitutes our problem.

2. The satisfaction serves equally well as a motive when a greater degree of reflective thinking is necessary in order to determine the method of response. The question "Did you break the glass?" has not become an adequate stimulus for the response "yes" or "no." The truthful answer depends upon a feeling of value associated with previous truthtelling experiences, an interest in the quality of the outcome; not upon an established connection through certain synapses.

In this type of conduct imagination functions significantly. Some image of the proposed response rises into consciousness; that image awakens the *feelings* which have *previously* been *associated with the outcome of that response.* The response is carried out or inhibited in terms of this affective product of imagination.

With motor skills—habits in the more exact sense—there is no perceptible stimulation of imagination. The stimulus drives directly to its appropriate motor neurones and the response comes about automatically. In these acts dependent on the worth of the outcome there is the momentary picturing of the result before the response is definitely aroused.

Further study of the functioning of consciousness in imagination promises to reveal more significance to this instantaneous image-patterning of the means toward desired ends than has yet been recognized.

For purposes of conduct guidance, whether it be for health, or safety or thrift or honesty or chivalry, we shall have to recognize that while specific motor responses can be cultivated in the very young baby the use of these responses becomes significant only after their

value has been established and recognized and has come to function affectively as the determiner of behavior.

This does not imply in any sense that we must abandon the reform of recent days and return to exhortation. It means rather that we must provide situations which will call for desirable behavior, or in some way set up such behavior; then *follow up the response with an affective tone which will give meaning to subsequent acts of the same general import in terms of that import*, not in terms of the sensori-motor arc. It means that ways of responding must find their significance in the result of the set, not in the blind completion of a fixed motor response. A young man who fired the furnace in a small college gymnasium had the habit of taking a plunge every morning immediately upon rising. Through a defect in the drain pipe the water was drained off one night. The boy responded in the morning in his habitual way, diving into the empty tank and crushing his head on the cement floor. This is not the sort of control society wishes to develop in matters of thrift or truth or honesty or generosity. Nor is it the only possible way, as we might be led to think. The enjoyment of savings, the thrill of vital health, the social approval of truth and honesty and altruism may be attached to acts until the value is recognized and held in readiness to respond to available image patterns as a feeling of interest or desire.

True, it is possible to cultivate these attitudes by an inductive rather than a deductive process. It is possible to make specific acts satisfying until a series of specific acts has been associated with a typical affective state, and thus all the past acts of the series, and all potential acts as well, are integrated into a type-act which thus becomes a trait or an ideal. It is quite unnecessary to begin with a broad, ill-defined ideal and attempt to inculcate it as an ideal before constituent responses have been made and integrated.

The fact of basic significance is that affective states are associated with outcomes. The extreme position recently defended that affective states may be identified with neural arcs "and therefore the feelings may be ignored" since they are but a function of habits must be altogether reversed both in the interests of consistent explanation and of practical guidance. Even if the feelings were to be identified as functions of synaptic resistance, they could not be ignored, but rather claim the greater attention. If the feelings are organically to be identified with habit then habit training must far more wisely and far more consistently utilize the feelings. And even if the feelings

are dependent on synaptic resistance they are not dependent on the degree of resistance in a fixed series of synapses. Rather they may be associated with synapses leading to a countless variety of motor neurones, any series of which may be stimulated, the connection depending on which is identified in imagination with the pleasure-producing image-pattern.

All in all, the distinction between "sensori-motor-arc-habits" and "satisfactory-outcome-habits" seems to conflict with no data claimed as basic by any school of psychology, while it does much to remove confusion in our descriptions and definitions of acquired behavior tendencies. It supplements the serious gap left by the attempt to classify all acquired tendencies as conditioned reflexes. And it prevents confusion in the field of conduct training which must inevitably result if trained teachers attempt to initiate and build up specific motor responses without establishing interests in terms of their outcome-values.

THE RELATION OF NERVOUS-EMOTIONAL STABILITY TO EDUCATIONAL ACHIEVEMENT

NOEL KEYS

University of California

AND

G. H. WHITESIDE

Jordan, N. Y.

The precise relationship between general emotionality and nervous instability on the one hand and intellectual and scholastic achievement on the other is a subject upon which psychologists have yet to agree. Clinical psychologists, and more particularly certain writers of the psycho-analytic school, have often lent countenance to the popular conception that emotionality and erratic temperament tend to go with superior intellectual endowments.[1] Parents of "nervous" and "sensitive" children are encouraged to console themselves with the reflection that it is from the ranks of such that our geniuses are drawn.[2] It is even suggested that emotional stability may act as a positive deterrent to achievement in scholastic pursuits. Thus, basing his argument upon results from the Colgate Mental Hygiene Tests given the men in two successive classes of Colgate University, Young[3] writes, "From the armchair, one would except the students who were psychoneurotic to be especially likely to fail in their college work. As a matter of fact, the exact opposite is revealed by our records. There are twice as many students with stable emotions leaving because of failure than students with unstable emotional outlets." And this despite the fact that "there is a zero relationship at college level between intelligence and psychoneurotic traits."

Educational psychologists, on the contrary, have tended to stress the negative correlation commonly found between intelligence and measures of nervous and emotional instability, and to infer that psychoneurotic traits must act as a serious handicap in school work.

[1] Burrow, Trigant: Permutations within the Sphere of Consciousness. *Journal of Abnormal Psychology*, Vol. XI, Aug.–Sept., 1916, pp. 178–188.

[2] Cameron, Hector C.: "The Nervous Child." London, 1923, p. 10. Lay, Wilfrid: "The Child's Unconscious Mind." 1919, pp. 60–69. Adler, Alfred: "Organ Inferiority and Its Psychical Compensation." 1917. Also" Understanding Human Nature." 1928.

[3] Young, J. Bateman: How Emotional Traits Predispose to College Failure. *Journal of Educational Psychology*, Vol. XVIII, Dec., 1927, pp. 631–636.

Among those who adopt this attitude might be mentioned Burt,[1] Gates,[2] Hollingworth,[3] Terman,[4] and Thorndike.[5]

The most carefully controlled investigation made in this field to date is perhaps that of Rosen[6] who in 1924 paired fifty children diagnosed as neurotic with fifty presumably normal children of like sex, nationality, age, grade, and school. Efforts to discover characteristics in the test performance of the neurotics which might help to detect and distinguish such children from their normal mates failed completely. No appreciable or significant differences between the groups were found in their performance on the several portions of the Pintner Educational Survey Tests, the Pintner-Paterson Performance Tests, the Thorndike Handwriting Test, four cancellation tests, or the Stanford-Binet. For example, no items could be discovered in the Stanford-Binet on which the neurotic group tended consistently to fail or to succeed as compared with the non-neurotic, the amount of "scatter" in the two groups was the same, and the IQ's obtained on a retest after an interval of two weeks were as constant for one group as for the other.

Even the average IQ's of the two groups were found to be closely similar, and the total educational scores almost identically the same. That this last was, however, in large part a result of the method of pairing by age and grade, became evident when the two groups were compared with a truly unselected sample of the school population. By this standard, both the neurotics and their "normal" mates were found to be decidedly "below average in age-grade status and below par in intelligence." The average Binet IQ was, indeed, but 92.

Nature of the Present Experiment

The present study may be regarded as, in a sense, a continuation and a check upon this latter phase of Rosen's investigation. It

[1] Burt, Cyril: The Unstable Child. *Child Study Magazine*, Vol. X, pp. 61–79. Oct., 1917.

[2] Gates, A. I.: "Psychology for Students of Education." 1923, Chaps. VIII and IX.

[3] Hollingworth, L. S.: "Special Talents and Defects." 1923, p. 70. "Gifted Children." 1926, pp. 126–131.

[4] Terman, L. M.: "Genetic Studies of Genius." Vol. I, 1925, pp. 500–506.

[5] Thorndike, E. L.: "Educational Psychology." Vol. III, 1913, p. 363.

[6] Rosen, E. K.: "A Comparison of the Intellectual Status of Neurotic and Normal Children in Public Schools." Teachers College, Columbia University Contributions to Education, No. 188, 1925.

endeavors in a different setting and by somewhat different techniques to throw further light upon the following questions:

1. What, if any, are the differences in intelligence as measured by the Stanford-Binet test, between children characterized by undue emotionality and an excess of psychoneurotic symptoms and children of the same age and school who are markedly stable in temperament?
2. What differences in school standing and educational achievement are observable between these two contrasting groups?
3. Given equal intelligence as measured by the Stanford-Binet, does the educational and scholastic achievement of the nervous or emotional child differ markedly from that of his more stable fellows; and if so, how?

The experiment as conducted was confined to Grades VI, VII and VIII of the public schools of Jordan and Weedsport in central New York state. The type of population in these two neighboring villages is so similar and the educational facilities and standards so closely comparable that for practicable purposes the two schools might be regarded as one. In all, seventy-eight children were involved from the Jordan school and one hundred four from the Weedsport, a total of one hundred eighty-two.

The criterion of nervous-emotional stability employed was not, as in the Rosen study, the pronouncement of a single school pscyhiatrist, but a composite of teacher's ratings and the pupils' own scores on the Woodworth-Cady Questionnaire.

The latter, a modification of the Woodworth Questionnaire of Psychoneurotic Symptoms, is too well-known to require detailed description here.[1] It consists of eighty-five questions which the child is requested to answer "yes" or "no." The following are typical examples:

Do you usually sleep well?	Yes	No
Are you afraid in the dark?	Yes	No
Have you often been punished unjustly?	Yes	No
Do you sometimes worry about your health?	Yes	No
Do you often have a hard time making up your mind?	Yes	No

Each of these questions was selected with a view to revealing some trait of character regarded by psychiatrists as symptomatic of psychoneuroticism; and the subject's "score" is simply the number of these unfavorable symptoms to which he confesses. The test as a whole has been found to possess a certain practical utility, and the reliability is

[1] For further particulars, see Terman, E. L.: "Genetic Studies of Genius." Vol. I, pp. 500–506.

good. On a recent reapplication of the questionnaire to several hundred children after a five-year interval, Terman obtained a correlation with their previous performance of no less than .60.

For the teachers' ratings, the writers employed a simplified scale of their own devising, as follows:

Name of Student Rated.....Teacher.....

Rating should be made independently, without comparison with other teachers. Make a check at any point on the line which best indicates your estimate of the student in the trait in question.

1. Amount of Laughter	(10) Near hysterics	(8)	(6) Giggles a great deal	(4)	(2) Self-controlled
2. Tendency to Moodiness	(2) Very stable	(4)	(6) Occasional fits of anger, blues or sulkiness	(8)	(10) Frequent and extreme excitement or despondency
3. Timidity and Fearfulness	(10) Very timid and easily frightened	(8)	(6) Very easily embarrassed. Sensitive to praise or blame	(4)	(2) Good courage and self-assurance
4. Restlessness	(1) Quiet and steady	(2)	(3) Wriggles about. Cannot continue long at one task	(4)	(5) Nervous twitching and spasms
5. Control of Speech	(5) Stammers a great deal	(4)	(3) Excited, jerky speech	(2)	(1) Calm, even speech
6. Tendency to Daydreams	(1) Very wide-awake and alert	(2)	(3) Mind wanders occasionally	(4)	(5) Lost in dreams most of the time

The six headings or counts on which each pupil was rated were selected as being those most likely to be brought to a teacher's attention. The numerals in parenthesis did not appear upon the forms submitted to the teachers, but are shown here to indicate the numerical value given to a check placed in any given section of the line. As will be noted, the first three points (amount of laughter, tendency to moodiness, and timidity or fearfulness) are given just twice the weight of the last three. This was done because of their more direct and obvious bearing upon emotionality. The sum of the various ratings is added to yield a total numerical score. Thus, a pupil given the

most unfavorable possible mark of all six heads would receive a total rating of forty-five in emotionality; while a pupil at the opposite, or stable, extreme would rate nine.

The moderate size of the classes and the unusual constancy or permanence of pupil and teaching personnel in the schools concerned tended to insure a close acquaintance between teacher and pupils most favorable for purposes of rating. It was found possible to secure for each of one hundred eighty-two pupils the independent ratings of three teachers, each of whom had known and taught the pupil for at least one year. These were in every case:

1. The pupil's regular classroom teacher during the current year. Her rating was given a weight of three in computing the average.
2. The pupil's regular teacher of the preceding year, whose rating was given a weight of two.
3. The teacher of special subjects (music and art), whose rating was given a weight of one.

The average of the three ratings thus obtained was accounted the "teachers' rating in emotionality" for the pupil concerned.

In order to permit of combining this average with the pupil's own score on the Woodworth-Cady Questionnaire, the ratings and the Woodworth-Cady scores were first converted into standard deviation scores. That is, the pupil's standing in each respect was expressed in terms of the number of standard deviations above or below the mean of the entire group. The two standard deviation scores were then averaged to obtain the pupil's composite score in nervous-emotional stability.

Pupils falling one or more standard deviations above the average of the group in terms of this composite criterion were designated the "Emotional Group" for purposes of the present experiment, and those lying one or more standard deviations below the average were accounted the "Stable Group." Each of these groups was found to number just thirty pupils.

Reliability of the Criterion

As a partial check upon both the reliability and the validity of the composite criterion described above, correlations were computed between the teachers' ratings in emotionality and the Woodworth-Cady scores. These are shown in Table I.

TABLE I.—CORRELATION OF PUPILS' SCORES ON THE WOODWORTH-CADY QUESTIONNAIRE WITH TEACHERS' RATINGS OF THESE PUPILS ON NERVOUS-EMOTIONAL STABILITY

SCHOOL	*N*	*r*	PE
Jordan	78	.40 ±	.062
Weedsport	104	.58 ±	.044
Combined group	182	.53 ±	.036

Having regard to the facts that the teachers' rating scale and the pupils' questionnaire differ somewhat in their emphasis and that no attempt has been made to correct for attenuation due to unreliability of the measures themselves, these inter-correlations are somewhat higher than might have been anticipated. It is evident that pupils' own answers to the Woodworth-Cady Questionnaire show a substantial correspondence with teachers' ratings of these pupils on the score of nervous-emotional stability. And it seems not unreasonable to regard this fact as a mutual confirmation of the validity of the measures themselves.

Furthermore, with what might be described as the two halves of our criterion displaying an inter-correlation of .53 ±.036, it is probable that the composite score would possess a reliability of .70 or better. The Spearman-Brown formula, which does not adequately cover such a case, predicts .69. For our present purpose, which is the establishment of two discrete or widely-differing groups, this must be regarded as entirely satisfactory. With a reliability of .69, the PE of an individual measurement is less than .38 of the sigma of the distribution. The difference between the lower limit of a group composed of pupils classed one or more sigma above average and the upper limit of a group falling one or more sigma below average is, therefore, approximately 5.3 times the PE of measurement. This means that the chance of any pupil properly belonging in the one group being wrongly classed within the other is less than 1 in 2000; and the chances are strongly against there being any overlapping between the two groups with respect to the true status of the individual pupils in nervous-emotional stability, as defined by our criterion.

COMPARISON OF THE EMOTIONAL AND STABLE GROUPS

A comparison of the thirty pupils selected as nervous and emotional with the thirty selected as most stable in these respects revealed no appreciable differences in the sex or national origins of the two groups. The number of boys and girls was the same in both groups, being

fifteen boys and fifteen girls in each. The racial composition was likewise almost identical. All of the children were white. A majority in each group were of British descent, and only five in each group were of Latin or Slavic stock.

A tabulation of the "scatter" or years spread above the basal year on the Stanford-Binet tended to confirm the findings of Rosen[1] and Mathews[2] as against the conclusions of Mateer[3] or Irwin and Marks,[4] in that the average number of years "scatter" was not appreciably greater for the emotional than for the stable group. The correlations between mental and educational age were likewise approximately the same for the two groups, being .69 for the emotional as compared with .65 for the stable.

In order to secure as accurate a measure as possible of the mentality and educational achievement of the two groups, the Standford-Binet Intelligence Test and the Stanford Achievement Test in full were given to each of the sixty pupils concerned. As in the case of the questionnaires, all of these tests were administered and scored by one of the present writers. The results are set forth in Table II.

Whereas no differences were discoverable in the sex or racial composition of the two groups, the differences in intellectual and educational status are great. These differences are, moreover, highly significant statistically, being, except in the case of chronological age, from four to seven times the probable error. *Children in Grades VI to VIII who are characterized by their teachers and by their own scores on the Woodworth-Cady Questionnaire as conspicuously nervous and emotional display a strong and reliable tendency to average more than one year retarded in age-grade standing,[5] nearly two years lower in mental and educational age, and some eighteen points lower in IQ as compared with their more stable classmates.*

These facts are undeniable. Their practical significance and implications as to the bearing of nervous and emotional traits upon scholarship are not quite so clear. Various circumstances may have operated to make the agreement between emotionality and inferior mentality

[1] *Op. cit.*, pp. 22–25.

[2] Mathews, Julia: Irregularity in Intelligence Tests of Delinquents. *Journal of Delinquency*, Vol. VI, March, 1921, pp. 355–61.

[3] Mateer, Florence: "The Unstable Child." 1924, p. 70.

[4] Irwin, E. A., and Marks, L. A.: "Fitting the School to the Child." 1924, p. 187.

[5] Table II shows the emotional group to be at the same time 6.8 months older and more than half a grade lower in school than the stable group.

TABLE II.—COMPARISON OF THE GRADE PLACEMENT, CHRONOLOGICAL AGE, MENTAL AGE, EDUCATIONAL AGE, AND INTELLIGENCE QUOTIENT OF PUPILS RANKING ONE OR MORE SIGMA ABOVE AVERAGE WITH THOSE PUPILS RANKING ONE OR MORE SIGMA BELOW AVERAGE IN NERVOUS AND EMOTIONAL TRAITS
(N = 30 in Each Group)

	Mean	SD	Difference (emotional-stable)	$PE_{Diff.}$	$\frac{Diff.}{PE_{Diff.}}$
Grade placement:					
Emotional group	7.67 ± .088	0.72			
Stable group	8.23 ± .096	0.79	− 0.56	0.13	−4.31
Age (in months):					
Emotional group	165.7 ± 1.9	15.2			
Stable group	158.9 ± 1.5	11.9	6.8	2.39	2.85
Mental age (in months):					
Emotional group	154.4 ± 2.6	21.2			
Stable group	177.6 ± 2.5	20.1	−23.2	3.59	−6.46
Educational age (in months):					
Emotional group	154.4 ± 2.4	19.4			
Stable group	177.9 ± 2.3	18.7	−23.5	3.32	−7.09
Intelligence quotient (Binet):					
Emotional group	93.9 ± 2.0	16.3			
Stable group	112.3 ± 1.6	12.8	−18.4	2.56	−7.19

appear larger than is in fact the case. It is altogether probable, for example, that the familiar "halo" illusion would cause certain teachers to rate as possessed of undesirable emotional traits those pupils whose general classroom work was poor. Insofar as this was the case, a high rating in emotionality would be not so much an indication of the true nervous and emotional traits of the child, as a reflection of the teacher's general impression of the pupil and his attainments. It is possible also that certain subtle factors of insight and self-consciousness might lead brighter pupils to confess to fewer of these evidently undesirable characteristics than their duller companions. Finally, it is not unlikely that the nervous and emotional child is at a certain disadvantage in test situations, so that his test scores, as such, do less than justice to his actual intelligence and achievement. Each and all of these factors would tend to produce a spuriously large correlation between emotionality and mentality in our present study.

TABLE III.—COMPARATIVE EDUCATIONAL ACHIEVEMENT OF TWO GROUPS MATCHED AS TO SEX, CHRONOLOGICAL AGE, MENTAL AGE, AND INTELLIGENCE QUOTIENT, BUT DIFFERING MARKEDLY IN NERVOUS AND EMOTIONAL TRAITS

Emotional Group

Pupil	Grade[1]	Age in months	Mental age	Intelligence quotient	Educational age (in months)
Boys					
BEJ	6	155	142	92	119
BUD	7	156	161	103	166
EDB	6	149	181	121	193
ESA	7	171	168	98	159
HOJ	8	187	181	97	181
KYK	6	157	150	96	129
MAW	7	162	153	94	155
MEW	7	143	194	136	193
Girls					
BAB	6	140	156	111	148
BRA	6	145	175	121	170
DAW	6	143	138	97	148
ELH	7	155	169	109	181
FLM	8	175	180	103	161
HAR	8	182	190	104	168
HAG	6	196	160	82	143
ROB	8	175	206	118	182
Mean.....	6.81[1]	161.9	169.0	105.1	162.3

Stable Group

Boys					
BOE	8	159	165	104	176
COJ	6	136	132	97	132
DOR	7	157	198	126	210
FAG	8	167	174	104	187
MIM	8	156	146	94	172
RAK	7	156	189	121	177
ROE	8	161	205	128	208
YAA	6	139	152	109	161
Girls					
COH	7	156	154	99	168
DOJ	8	161	195	121	184
GOV	8	165	167	101	176
HOR	8	175	168	96	190
LDO	8	187	161	86	177
VDV	7	190	181	95	164
WAA	7	145	148	102	155
ZEE	8	165	183	110	182
Mean.....	7.44[1]	160.9	169.9	105.8	176.2

[1] The grade as shown in this column is merely the number by which the grade is known. Since the tests were given in May, when the work of that grade was nine-tenths completed, the true grade standing of the groups would average approximately 0.9 higher than the means here shown. Tables II and IV give means as corrected.

In order to determine as accurately as possible the actual influence of nervous-emotional traits upon scholastic achievement, it becomes necessary to minimize or eliminate the obscuring effects of the irrelevant factors mentioned. To do this, the writers sought to secure two groups, matched as to age and intelligence test scores, but differing widely in their nervous-emotional characteristics. The educational achievement of these two groups could then be compared without prejudice.

The results are set forth in Table III and summarized in Table IV. Owing to the marked differences already noted in the age and intelligence of the emotional and the stable groups, it proved impossible to include more than sixteen of the thirty pupils in each of the final matched groups. While the average age and IQ of the resultant groups are not absolutely identical, the differences, as will be seen from Table IV, are too small for either practical or statistical significance. Differences in grade standing and educational attainment as measured by the Stanford Achievement Test are, however, large. The emo-

TABLE IV.—COMPARISON OF THE GRADE STANDING AND EDUCATIONAL AGE OF TWO GROUPS MATCHED AS TO SEX, CHRONOLOGICAL AGE, MENTAL AGE, AND INTELLIGENCE QUOTIENT
(N = 16 Each)

	Mean	SD	Difference (emotional-stable)	$PE_{Diff.}$	$\frac{Diff.}{PE_{Diff.}}$
Grade placement:					
Emotional group	7.71 ± 0.15	.88			
Stable group	8.34 ± 0.12	.70	− 0.63	.19	−3.3
Age (in months):					
Emotional group	161.9 ± 2.7	16.6			
Stable group	160.9 ± 1.7	10.1	1.0	3.19	0.3
Mental age (in months):					
Emotional group	169.0 ± 3.1	18.5			
Stable group	169.9 ± 3.4	20.2	− 0.9	4.60	−0.2
Intelligence quotient (Binet):					
Emotional group	105.1 ± 2.2	13.2			
Stable group	105.8 ± 2.0	12.0	− 0.7	2.99	−0.2
Educational age (in months):					
Emotional group	162.3 ± 3.5	20.8			
Stable group	176.2 ± 3.0	17.8	−13.9	4.61	−3.0

tionally unstable group are, on average, .63 of a year behind in grade placement and more than a year (13.9 months) lower in educational age as measured by test performance. Despite unreliability due to the smallness of the numbers, these differences are sufficiently great to possess considerable statistical significance, being more than three times the probable error. In other words, the chances are 98 in 100 that a repetition of this experiment under like conditions would show the nervous-emotional children inferior to their stable fellows in grade standing and educational achievement.

Dangers in the Use of the Accomplishment Quotient

A comparison of Table II with Table IV serves incidentally to bring out most forcefully the fallacies likely to result from use of the accomplishment quotient as a measure of relative achievement. Table II shows the thirty pupils in the emotional group to average almost identically the same in educational as in mental age. The same is true of the stable group. The figures are 154.4 against 154.4, and 177.9 against 177.6. The mean accomplishment quotient is, therefore, 100 for each group. The implication is clearly that the emotional and the stable groups have done equally well in their school work in proportion to their intelligence; that each has, in fact, accomplished exactly what was to be expected.

It must be remembered, however, that these two original groups differed markedly in age and tested mentality. The "stable" group was the younger and brighter. The question is whether the AQ affords a just and adequate basis of comparison in such cases.

From the time of its introduction, the AQ has been found to correlate negatively with IQ.[1] The effect is to make dull pupils appear accelerated and bright pupils retarded in educational achievement. But, as certain recent writers have shown,[2] this phenomenon is an

[1] Franzen, Raymond: "The Accomplishment Ratio." New York: Teachers College, Columbia University, Contributions to Education, No. 125, 1922, p. 14. Also Ruch, G. M.: The Achievement Quotient Technique. *Journal of Educational Psychology*, Vol. XIV, September, 1923, pp. 334–343.

[2] Wilson, W. R.: The Misleading Accomplishment Quotient. *Journal of Educational Research*, Vol. XVIII, January, 1928, pp. 1–10. Douglass, Harl R. and Huffaker, C. L.: Correlation between Intelligence Quotient and Accomplishment Quotient. *Journal of Applied Psychology*, Vol. XIII, February, 1929, pp. 76–80. Cox, Genevieve L.: A Study of Various Factors Which Influence the Use of the Accomplishment Quotient. *Journal of Educational Research*, Vol. XXI, January, 1930, pp. 29–42.

almost inevitable result of the AQ technique, which, in itself, gives an undue advantage to the dull. Errors in the measurement of intelligence tend to exaggerate the mentality of the bright and the intellectual handicaps of the dull. Moreover, achievement is conditioned not only by mentality but by health, environment, and emotional traits. Since these correlate only moderately with intelligence, it is unreasonable to demand that school achievement be strictly proportional to the latter. Finally, it must be recognized that, with the methods of promotion now in vogue, slow pupils are almost inevitably accelerated in grade placement in proportion to their mental age, and bright pupils as inevitably retarded. The superior quickness of the latter cannot entirely offset the advantage which more advanced instruction has given to the slow. All things considered, it is a much simpler matter for the mentally sub-average child to achieve a high AQ than it is for his more gifted brother.

It follows that any attempt to compare the achievement of children of widely different ages and IQ's on the basis of AQ alone is almost certain to distort the facts. In the present study, when emotional children of an average IQ of 94 are compared with stable children of an average IQ of 112, as in Table II, the mean AQ's of the two groups are found to be the same. When, however, the brighter of the emotional pupils are matched against the duller of the stable pupils so as to obtain two groups *of the same average age and intelligence,* as in Table IV, a striking difference in their achievement becomes apparent. The average educational age of the stable pupils in Table IV is fourteen months higher than that of the emotional; and the mean AQ, *i.e.*, EA/MA, of the stable group is 103.7 as against only 95.4 for the emotional.

Summary and Conclusions

1. There is no tendency for nervous and emotional instability to go with superior intelligence and knowledge among children of junior high school age. On the contrary, there is a well-marked and reliable tendency for pupils characterized as conspicuously nervous and emotional to fall far below the emotionally stable in intelligence and school standing. In the Grades VI, VII and VIII here studied, the nervous and emotional children averaged one year retarded in age-grade status, nearly two months lower in mental and educational age, and eighteen points lower in IQ than did their markedly stable classmates.

2. When children of like sex, age, intelligence, and nationality, but widely different nervous and emotional traits are compared, the nervous and emotional tend to average appreciably lower in grade placement and educational achievement than the emotionally stable. In the present study the accomplishment quotient of the emotional averaged 95.4 as contrasted with 103.7 for the stable.

3. Children's scores on the Woodworth-Cady Questionnaire of Psychoneurotic Symptoms correlate well (in this study, .53 $\pm$ .036) with an average of three teachers' ratings of these children on nervous and emotional traits.

4. The accomplishment quotient should not be used to compare the relative educational achievement in groups which differ widely in chronological age and IQ.

THE BIOLOGICAL SIGNIFICANCE OF INTELLIGENCE TESTS

DAVID W. OATES

Municipal Secondary School for Boys, Head Master, Newport, England

Introduction

As a further development of her work on Will-temperament, June E. Downey has been assisted by Richard S. Uhrbrock in an effort to devise non-verbal material that would induce the same mental sets as are obtained by the use of the verbal form of test. The non-verbal tests[1] are not yet available for general distribution, but in an investigation conducted by the writer with a copy of the test supplied by the authors some points of interest have arisen. One or two of the tests produced results which showed no apparent age increment in the various age groups and this appeared to be a matter worth further investigation. The results here presented were obtained with 286 boys of eleven to eighteen years of age in attendance at a Secondary School.

The Tests Used

The main tests used in this investigation are those named by the authors "Speed of Movement" and "Ability to Speed." The test material consists of a prepared sheet of paper on which are two sets of horizontal lines 16.5 cm. long and 8 mm. wide. At the left-hand end of the first pair of horizontal lines are five parallel slanting lines followed by a small arrow. The directions for the Speed of Movement test used in this investigation were as follows: "Notice the slanting lines between the two long lines. Also, notice the arrow. When I say 'Start,' begin at the arrow and draw lines parallel to the slanting lines, and the same distance apart, until I say 'Stop.' Work at your natural speed. Ready. Start." For the second part of the test designed to measure Ability to Speed, the following additional instructions were used: "This time you are expected to work as fast as you possibly can. When I say 'Start,' draw as many lines as you can in the time allowed. Work rapidly. Ready. Start. Fast." The time allowed in each case was twenty seconds. The results which are given below were obtained by calculating for each subject the ratio between the

[1] Uhrbrock, Richard S. and Downey, June E.: A Non-verbal Will-temperament Test. *Journal of Applied Psychology*, Vol. XI, No. 2, April, 1927, pp. 95–105.

number of slanting lines made in the second test and the number of slanting lines made in the first test.

For purposes of comparison and analysis, the subjects were also given a series of Intelligence Tests which included Cipher, Absurdities, Analogies, Number Series, and Best Reason Tests. The results of Ballard's Crichton Test obtained in another investigation are also given.

The Results

Before considering the results of the tests it was considered desirable to determine how far whatever qualities are measured by these Ability to Speed Tests are constant, and how far they are affected by changes in the nature of the test itself. For this purpose extended tests were conducted with a group of sixty-one boys. The test in its original form was repeated. Then variation was introduced first by making the lines slant in the opposite direction, then by substituting the repeated writing of the subject's own name, and the continuous writing of the figures one to nine in place of the slanting lines. The following results were obtained for which full details need not be given: The average ratios for the group were, Original test 1.64, reversed slope 1.56, Writing name 1.58, writing figure 1.59: thus to one place of decimals the four methods of test result in a ratio of 1.6. The mean variation was very similar in each case and the correlation between the results was high. We are, therefore, apparently tapping some fundamental quality in the boys forming this group.

The results of the different tests in the various age groups are given in the following table.

Table I

Age	*N*	Ability to speed	Mean variation	Intelligence tests	Mean variation	Crichton test	Correlation of ability to speed and intelligence tests
11 to 11-11	16	1.72	.30	61.4	6.1	....	−.216
12 to 12-11	62	1.67	.39	65.2	7.2	8.8	.065
13 to 13-11	68	1.75	.54	68.5	8.1	13.7	−.523
14 to 14-11	63	1.58	.32	78.0	7.5	17.2	−.178
15 to 15-11	51	1.79	.41	83.8	6.1	18.3	.223
16 to 16-11	14	1.52	.29	84.2	4.9	25.75	−.103
17 to 17-11	7	1.65	.26	88.4	4.1	29.4	−.050
18 to 18-11	5	1.47	.10	85.9	3.6	31.0	−.622

While there is a gradual increase in score in successive age groups in the intelligence tests, we find no corresponding change in the scores for Ability to Speed as measured by this particular test. The mean variation is lower in the older age groups, but this is probably due to the smallness of the groups tested. The correlation between Ability to Speed and Intelligence is definitely negative with the exception of the coefficient for the fifteen year group and in this case the PE is .0897 so that the coefficient is less than three times its PE. The average coefficient of correlation for Ability to Speed and Intelligence is $-.175$. The coefficient of correlation with age is $-.595$, so that the particular quality has no apparent relationship to age in the age groups forming the subjects of this investigation.

The relative results of the tests are more clearly seen in the following graph. In order to make it possible to establish a definite standard of comparison between the separate tests the scores in each age group

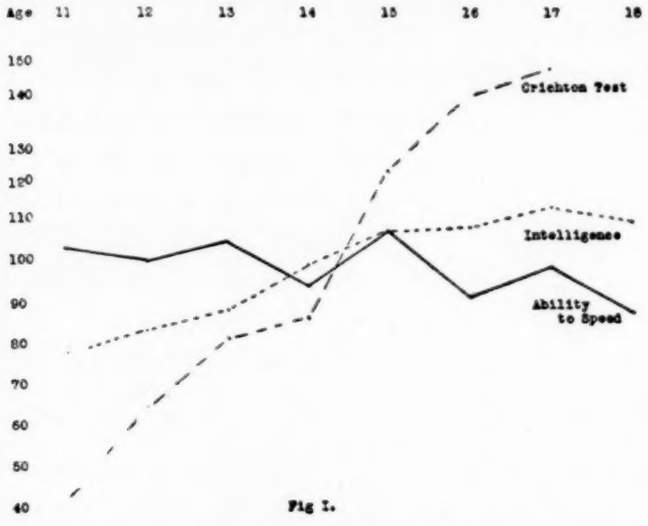

Fig 1.

in a particular test have been expressed as a per cent of the average score for all the subjects in that particular test.

No results can be quoted for comparison, as the authors of the test, so far as can be gathered from their published report, have not obtained age norms for these particular tests. It is interesting to find, however, that the scores obtained by Downey for the Freedom from Load test which is an index of inertia as revealed in the ability to maintain speed without external pressure and is based upon the ratio of speeded to normal time for writing, fully supports our results with this test. The results[1] were as follows.

[1] Downey, J. E.: "The Will-temperament and Its Testing." 1923, p. 253.

TABLE II

AGE	*N*	SCORE
12	48	133
13	53	143
14	66	133
15	39	132
16	15	139
17	20	131
18	34	127

Downey also obtained with a group of forty subjects aged 21–25 an average score for this test of one hundred thirty. These results

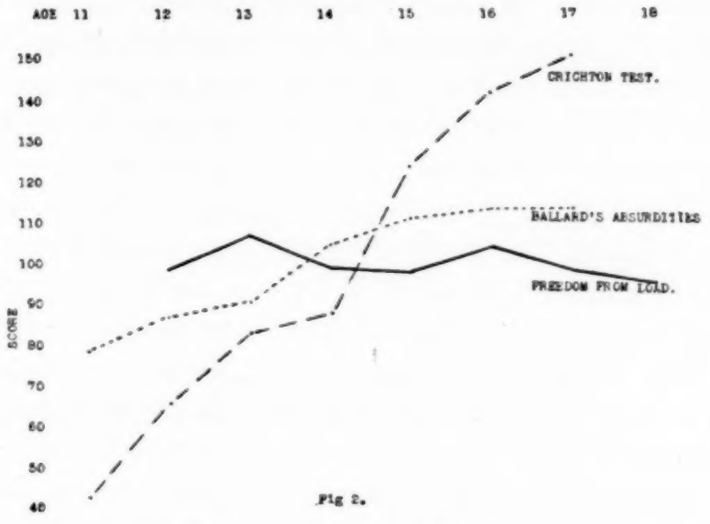

Fig 2.

are given in Fig. 2, together with the intelligence scores for the corresponding age groups calculated as indicated above, and the scores obtained by Dr. Ballard using the Absurdities test.

DISCUSSION OF RESULTS

Whether we can regard the average number of points scored by the subjects in each age group in a particular test as an intelligence growth curve is a question that must be handled with caution. We have been reminded[1] that different conclusions regarding the rate of mental growth might be arrived at from the same data according to whether we calculate the amount of work as expressed in score points, done in unit time, or whether we calculate the amount of time required

[1] Peterson, J.: The Growth of Intelligence and the Intelligence Quotient. *Journal Educational Psychology*, Vol. XII, No. 3, March, 1921, pp. 148–154.

for performing a unit of work such as a score point in the scale. Here, however, we are not concerned with reading actual mental growth from year to year directly from our curves, but are using the curves as approximate age progress curves expressed in terms of annual increments for purposes of comparison with age progress curves obtained by other tests. This theoretical consideration will not, therefore, affect the validity of our conclusions. The curves for intelligence are in close agreement with those obtained in other investigations. Burt[1] using reasoning tests, and Ballard[2] using Absurdities test have produced curves indicating that mental growth is fairly constant up to fourteen years of age, but that in the fifteenth and sixteenth years it grows less and then ceases. In the Ability to Speed test and the closely related Freedom from Load test, however, there is no significant change from one year to the next, that is, there are apparently no annual increments of growth; the eleven year old subjects are as efficient in these qualities as the eighteen year old subjects. It appears that whereas the mental growth curve reaches its flat point about the age of sixteen years, the particular efficiency tested by the Ability to Speed test has already reached its highest point at or before the age of our youngest group of subjects. It is interesting to note in this connection that Childs[3] investigating ability in Drawing with 2177 children found that all appreciable growth of improvement ceased at nine or ten years of age.

Some of the results of other investigations have a bearing upon the problem we are considering. Stead[4] found that boys of twelve to thirteen on the average scored better than those of thirteen to fourteen or fourteen to fifteen in all motor tests except those for steadiness, and that ability in motor and mental activities "hang more closely together in the case of younger children than they do in the case of older ones." Muscio[5] on the other hand found no correlation between mental and motor activities in adults. Farmer[6] also has pointed out that "fairly high correlations have been found between

[1] *British Journal of Psychology*, Vol. XII, 1921. Cf. Burt, C.: "The Curve of Mental Growth," "The Measurement of Mental Capacities." 1927, p. 26.

[2] *Journal of Experimental Pedagogy*, Vol. V, 1919.

[3] *Journal of Educational Psychology*, Vol. VI, 1915.

[4] Stead, H. G.: Factors in Mental and Scholastic Ability. *British Journal of Psychology*, Vol. XVI, No. 3, 1926, pp. 203, 204.

[5] *British Journal of Psychology*, Vol. XIII, No. 2, 1922, p. 184.

[6] Farmer, E.: A Group Factor in Sensory Motor Tests. *British Journal of Psychology*, Vol. XVII, No. 4, 1927, p. 331.

certain motor tests and tests of intelligence, but that these intercorrelations tend to become smaller as the age of the children increases" and suggests the explanation that motor tests for young children are really tests of intelligence inasmuch as among children with partially developed intelligence it is only the most intelligent who understand what is required in a motor test. This explanation suggests a weakness in administering or designing the particular tests; other explanations such as the one we develop below might also be urged.

Our results it is suggested are satisfactorily explained on the biological theory of development as representing successive stages in maturation of function. In the motor reactions called for in the Ability to Speed test we are dealing with movements of a type which are amongst the earliest of our inherited capacities. "Natural development . . . keeps on until maturity. The neurones continue to grow, and their synapses in the nerve centers to become closer knit, just by virtue of natural growth; and thus reflex arcs, and other reaction machinery, one by one reach the ready-to-use stage during the individual's growing up."[1] There is some confirmation of this in a paper recently submitted to the American Psychological Association by Travis[2] who reports an increase in the rate of conduction of the nervous impulse with increasing age, and suggests that neurophysiologically we may attribute intelligent behavior to the relatively simultaneous interactivity of the higher levels of the central nervous system.

The development of the mental life of an individual may be traced in the gradual growth to functional maturity of the various nerve centers. Some attain their full development and functional maturity rapidly; others develop slowly through easy stages, and the period of growth to functional maturity is more prolonged. By means of correlations and coordinations the individual progresses from simple to broader and more complex forms of activity, and the marvellous complexity of adult intelligence is but the enrichment and refinement of powers that emerged at relatively early stages of development. Spearman[3] explains all cognitive growth as consisting in a progressive clarification. "This clarification may be likened to the ripening of a fruit . . . Tests of *g* should not be limited to

[1] Woodworth, R. S.: "Psychology." Seventh Edition, 1927, p. 92.

[2] Travis, L. E.: A Neurophysiological Envisagement of Intelligence. *Psychological Bulletin*, Vol. XXVI, March, 1929, p. 145.

[3] Spearman, C.: "The Abilities of Man." p. 216.

ascertaining whether the fruit has already fallen; they should measure instead the degree of ripeness at whatever stage is most appropriate to the ability of the persons tested."

We have in the data before us a fairly complex function in our Ability to Speed test which apparently has reached maturity before the youngest age group were tested; that is, the fruit has already fallen, and on the other hand, in the Intelligence tests, biologically later functions which are developing through the age groups revealing increasing degrees of ripeness and reach maturity in the oldest age groups under investigation. Farmer[1] approaching this problem from a different point of view arrives at a similar conclusion. He found no correlation between general motor abilities and a correlation of about .25 for æstheto-kinetic tests involving a high degree of coordination and mental control for their performance. "This degree of correlation," he remarks, "is intermediate between that of motor tests in general, which do not inter-correlate at all, and intelligence tests which inter-correlate highly. Coordinated movement requiring conscious direction is biologically intermediate between the ability merely to move and the ability to think." He then points out that these functions are affected in reverse order under the influence of an anæsthetic and concludes: "This parallelism between the magnitude of the average inter-correlation coefficients and the biological level of the functions from which they are obtained is curious, and if it turned out to be a true relationship, it might lead to a useful classification of psychological tests."

Spearman recognizes the exceptionally low correlation existing between motor abilities and explains it on the assumption that "these operations bring largely into use the peripheral parts of the nervous system, and even non-neural structures." "These [biologically earlier operations] do not subserve mental processes [biologically later functions] directly, but only influence them indirectly."[2]

Our data can be regarded in another way. Black has shown that it is "theoretically possible that in connection with any instinct or engrain-complex (or group of either) the appreciation of meaning mechanism and the reaction mechanism might have distinct efficiencies . . . A test of reaction or 'adaptability' gives, therefore, a coefficient of efficiency which is compounded of the two coefficients of efficiency

---

[1] *Op. cit.*, p. 330.

[2] *Op. cit.*, p. 217. The material within the brackets is mine. D. W. O.

(1) of appreciation, (2) of reaction."[1] We may regard our Ability to Speed as biologically the earlier function and consisting mainly of the efficiency of reaction and only to a very limited extent of efficiency of meaning. We have, for example, in the gradation from simple reaction, through choice reaction to associative reaction a series of functions each of which calls for a longer time factor for its performance and may be regarded as involving an increasingly larger element of meaning. Koffka is referring to this similar fact of development when he says: "Whereas originally the reaction follows directly upon the stimulus, intervening members become more and more numerous, and more and more important, as development progresses . . . What we call mental work is for the most part work done with these intervening members."[2] The biological significance of these facts is evident on the assumption that efficiency of reaction reaches functional maturity earlier in the life of an individual than efficiency of meaning does.

This may throw some light upon an interesting feature in our results. The curve for the battery of intelligence tests is intermediate between that for Ability to Speed and for the Crichton test. This may be loosely attributed to the fact that the Crichton Test is too difficult for the subjects in the earlier age groups, while some of the tests in the battery of tests are too easy for the subjects in the older age groups. We may now examine a little more closely the significance of the terms "too difficult" and "too easy," and suggest that a test is too easy when it introduces too large a proportion of operations that have already reached maturity at the given age, and is too difficult when it involves too large a proportion of operations that are, at the given age, too far removed from maturity. Similarly, we can explain the curves in our graph by assuming that the Ability to Speed test is largely a test involving efficiency of reaction, while the tests constituting our battery of tests involve a much larger proportion of efficiency of meaning and the Crichton test involves almost entirely efficiency of meaning. The last-named will, therefore, form the best test of intelligence for those ages in which intelligence is reaching maturation. This is what Spearman means when he writes: "Only when the critical processes in the tests are restricted to this eduction is there any hope of really determining the point where *g* attains to

[1] Black, C.: Note on the Nature of Intelligence. *British Journal of Psychology*, Vol. XVIII, No. 4, April, 1928, pp. 453–454.

[2] Koffka, K.: "The Growth of the Mind." 1914, pp. 148–149.

its maturity."[1] His contention is that the problem of determining the upward growth of intelligence will be really insoluble "until the distinction is introduced between retention on the one hand and eduction on the other," only the latter of which he holds is involved in *g*. Similar in nature is Black's differentiation between efficiency of reaction and efficiency of meaning. Does intelligence as far as we understand it involve only the latter?

The score in an intelligence test is always a measure of a complex thing. Differences in performance may be ultimately attributed to the effect of a number of variables, such as heredity, maturity, the unorganized environmental and nurtural influences of the home, and the more highly organized influences of school training. Fundamental powers have developed primarily as a biological equipment to meet simple recurrent and urgent situations, and equally obvious is the fact that biologically from infancy to manhood there is an urge compelling man to engage in different kinds of activity. A consideration of the data here presented emphasizes the necessity for analyzing the activities involved in intelligence tests from the point of view of the functional maturity of the processes underlying them. The nature of the curve of growth of intelligence may be determined largely by the appropriateness of the test material to the functional maturity of the processes to be expected at a particular age level. The testing conditions are now sufficiently under control to assume that the scores yield a reliable measure of ability, and sufficient refinements have been developed to keep nurture influences relatively constant so that a score may be assumed to be a fairly reliable measure of capacity. We may now, it is suggested, focus attention more definitely on the problem of maturation by studying more closely the functional levels of the various processes involved in our tests.

[1] *Op. cit.*, p. 362.

ADULT AGES AS A FACTOR IN LEARNING

HERBERT SORENSON

University of Minnesota

Learning ability probably follows very closely the mental growth curves. If one considers intelligence to be the capacity to learn, the mental growth curves and learning ability may be considered identical. Intelligence, as we conceive it, can not be separated from learning if mental growth is expressed by the amount learned, and future growth a function of what has been learned. According to such an hypothesis, individuals may vary in their intelligences according to their volitional traits for study as well as their opportunities; and likewise for learning.

The mental growth curve accelerates very rapidly up to the ages of ten and eleven when it rounds off to a plateau at the ages of fourteen, fifteen and sixteen. Heinis,[1] by statistical means, continues a slightly rising mental growth curve to an age of about forty. This suggests that capacity to learn increases with greatest rapidity in the pre-adolescent ages, and following those ages, increases slightly to forty. There are no curves experimentally determined for the complete span of life, but it is known empirically that there is a definite decline in late senescence. We do not know if intelligence, exclusive of the dementias, increases or decreases with advancing age. If we may use interchangeably the intelligence determined by intelligence tests and learning ability as obtained by investigation there is some ground for stating that intelligence does not decrease much with increasing age. Inferences that it does not decrease at all are even permitted.

At this point we may drop the concept of intelligence and limit ourselves to ability to learn. If the reader wishes to substitute intelligence as an equivalent for learning ability, the writer is not the one who will quarrel with him. However, the thesis which will be considered for major discussion is one concerning the relation of a wide range of adult ages to ability to learn.

In his experimentation reported in "Adult Education," Thorndike[2] concludes that there is a slight decrement in people's abilities to learn as they grow older. This decrease can be expressed quantitatively at about ten to fifteen per cent. In discussing the learning curve in Esperanto as determined by a class of adults Thorndike says:

[1] Heinis, H.: A Personal Constant. *Journal of Educational Psychology*, Vol. XVII, March, 1926, p. 167.

[2] Thorndike, E. L.: "Adult Education." Macmillan, 1928, 335 pp.

All evidence points to a curve for ability to learn a systematic logical language with a vocabulary based largely on Latin, French, German, English, and Italian, of the following type for individuals who attend school through college: Rising from eight to sixteen and probably to twenty, or later, then remaining parallel to its base line to twenty-five or later, then dropping very, very slowly to thirty-five, and somewhat more rapidly, but very slowly, to forty-five or later.[1]

Explanations for this decrease may be due to an inner degeneration of the nervous system, to disuse or long absence from systematic study, to lessened energy and enthusiasm with increasing age, or to lessened opportunity for study caused by heavier vocational demands.

Much of the experimental work connotes that the decreased learning ability manifested by older subjects is due to disuse. A kind of "rustiness" has set in due to the abandonment of vigorous study. Thorndike alludes to it as follows: "Adult learning is, itself, probably a partial preventive or cure for adult inability to learn."[2]

This paper proposes to provide evidence that such is the case. Effort will be directed to show that in a class made up of students who were resuming study after a long "lay off," age was a restrictive factor in learning, but that age was not a *limiting* factor in classes in which its students had done considerable university work previous to the class in which this study was made.

Three extra-mural classes or groups were included in this investigation. The classes consisted of elementary public school teachers who took courses offered for the Bachelor's degree in Education. Due to the nature of the courses, one was selected by teachers many of whom were resuming their formal study for credit and diploma after long absence from such pursuit. In the other two classes, the enrollments consisted of teachers who had recently taken several university courses.

Certain data were collected to provide information which would enable computations for ascertaining the effect of age on learning ability. Intelligence or college ability was obtained by the Miller Analogies Test; achievement by a very comprehensive objective examination; and amount of study by weekly report. It is with humility that one ventures to reveal how such an important item as amount of study was obtained. The method has no proved reliability nor is there any means of knowing its validity. The writer does think, however, that the results obtained by report were as reliable as such

[1] Thorndike, E. L.: *Ibid.*, p. 239.

[2] Thorndike, E. L.: *Ibid.*, p. 146.

data can be. They were carefully collected every week, coöperation was willing and thoughtful, and the class members understood perfectly as early as the first meeting of the class that their marks would be determined by their scores on the final examination independent of any reported amount of study or preparation.

Age distributions for the groups studied show ranges of thirty-six years with an average age of about thirty-five. Each age group is well represented except the very extremes.

TABLE I.—AGE DISTRIBUTIONS OF THREE CLASSES OR GROUPS

Ages	Group I	Group II	Group III	Total
56	1	..	..	1
52	2	1	1	4
48	3	4	1	8
44	4	4	2	10
40	5	2	5	12
36	5	8	12	25
32	10	7	8	25
28	18	7	11	36
24	7	8	10	25
20	3	..	4	7
Total	58	41	54	153
Average	35	36	33	35

To obtain evidence of the effect of age on learning or achievement the partial-correlations technique was used. It seems entirely justifiable to do so as none of the items is a cause or result of one or the other. They apparently are discrete and not integral functions of each other, either in causation or structure. This is particularly true of the age data and amounts of study.

Table II contains zero order correlations for the group which previous to this course had been relatively inactive in terms of systematic study. This group will be referred to as Group I while the other classes which were made up of students who previously had done considerable university study, will be designated as Group II and Group III.

There is really no relation between the age and the intelligence of this group of elementary school teachers. Those who are older tend to do more studying. In spite of that practice the r of age and achieve-

TABLE II.—ZERO ORDER CORRELATIONS OF SEVERAL VARIABLES FOR GROUP I (RESUMING STUDY)

	Age	Analogies[1]	Hours of study	Achievement[2]
Age	.....	−.10 ± .09	.40 ± .07	−.12 ± .09
Analogies	−.10 ± .09	.....	−.29 ± .08	.50 ± .07
Hours of study	.40 ± .07	−.29 ± .08	.....	.13 ± .09
Achievement	−.12 ± .09	.50 ± .07	.13 ± .09	

$N = 58$.

[1] An intelligence test designed for universty students. It has been used several years to test entrants into the University of Minnesota College of Education.

[2] Determined by a final objective examination containing over three hundred items.

ment is about zero. It appears that it doesn't pay to study, as the relation between hours of study and achievement is negligible. This relationship, however, is due to the failure of the brighter to study as much as those who score lower on the analogies exercise. The r of analogies and hours of study is $-.29 \pm .08$ while the r of analogies and achievement is $.50 \pm .07$.

In Table III are given some partial correlations for data found in Table II.

TABLE III.—PARTIAL CORRELATIONS OF MEASUREMENTS FOR GROUP I ("RUSTY" GROUP)

Correlation	Factor held constant	
First Order		
Age with achievement	Analogies	−.09
Age with achievement	Study (hours of)	−.19
Age with study	Analogies	.45
Achievement with analogies	Study (hours of)	.57
Achievement with study	Analogies	.39
Second Order		
Age with achievement	Analogies, study	−.32

The rather low correlations between achievement and mental ability with amount of study when one or the other in turn is held constant suggest that some qualitative factors are operating which

have been undetected. This may be particularly true of amount of study where multiple degrees of effectiveness are probable.

When age and achievement are related and analogies or amount of study held constant, there appears a slight inverse correlation. Following the suggestion of these negative correlations, both analogies and study were held constant in a partial of the second order and then the relationship between age and achivement becomes −.32. At this point the significant conclusion of the article can be made. There appears to be an age factor operating in this group. The inverse correlation between age and achievement is not strikingly high, but it does suggest that a slight disability for learning sets in with increasing age. It is probably due to a "rustiness" not only of the learning processes but of general orientation which when broken down to its parts includes note taking, effective study, adjustment to the objective examinations, ability or willingness to cram and other ways and means which contribute to achievement.

The evidence obtained from two other groups when contrasted with the first group tends to support Thorndike's supposition when he suggested that learning in itself might prevent decreased learning ability.

Table IV contains for Groups II and III (active groups) zero correlations corresponding to those given for Group I. One of the greatest difference in the correlations is found between age and achievement for the two groups. For Group I it is −.12 and for Group II it is .17. One must interpret such correlations by recognizing the limitations imposed by small r's and comparatively large probable errors.

Group III, with a slight positive correlation between age and college ability, with a zero correlation between age and hours of study as against .41, and a smaller negative correlation between college ability (analogies) and hours of study, differs a little from Group II. Group III differs from the other groups in the correlation of age with hours of study. Between analogies and hours of study the r's are negative for Groups I and II but practically zero for Group III. Excepting for this last group the older students had a penchant for more study while the bright study less. There is no significant relationship between age and mental ability.

In general, the relationship indicated for these classes are like those found by Jones and Ruch.[1] They also found that poorer students

[1] Jones, Lonzo and Ruch, G. M.: "Achievement as Affected by Amount of Time Spent in Study." Twenty-seventh Year Book. National Society for the Study of Education, Part II, p. 133.

TABLE IV.—ZERO ORDER CORRELATIONS OF SEVERAL VARIABLES FOR GROUPS II AND III (HAVE BEEN ACTIVE IN STUDY)

		Age		Analogies		Hours of study		Achievement	
		II	III	II	III	II	III	II	III
Age	II	.....	.....	−.02 ± .10	.....	.41 ± .09	.....	.17 ± .10	
	III	.....	.....	.....	.16 ± .09	.....	.00 ± .09	.....	−.03 ± .09
Analogies	II	−.02 ± .10	.....	.....	.....	−.32 ± .09	.....	.35 ± .09	
	III	.....	.16 ± .09	.....	.....	.....	.06 ± .09	.....	.48 ± .07
Hours of study	II	.41 ± .09	.....	−.32 ± .09	.....	.....	.....	.16 ± .10	
	III	.....	.00 ± .09	.....	−.06 ± .09	.....	.....	.....	−.06 ± .09
Achievement	II	.17 ± .10	.....	.35 ± .09	.....	.16 ± .10			
	III	.....	−.03 ± .09	.....	.48 ± .07	.....	−.06 ± .09		

study more than do the brighter. An inverse correlation of $-.28 \pm .055$ between hours of study and grade points (achievement) may be compared with .13, .16, and $-.06$ which when interpreted suggests negative relationship in the investigation of Jones and Ruch as against no correlation in the groups studied by the writer.

The partial correlations for the data of Table V are obtained from the correlations given in Table IV. In Table V, the partials that are most important to the thesis of this discussion are those of the

TABLE V.—PARTIAL CORRELATIONS OF VARIOUS FACTORS

Correlations	Factor held constant	Second group	Third group
Age with achievement	Analogies	.19	−.12
Age with achievement	Study	.11	−.03
Age with study	Analogies	.43	.02
Achievement with analogies	Study	.42	.48
Achievement with study	Analogies	.30	−.03
Age with achievement	Analogies, study	.07	−.12
N	.....	41	54

second order which concerns the relationship of age and achievement when two very important variables, ability and amount of study, are held constant. The .07 for Group I and $-.12$ for Group III is very slight and insignificant. They are not as large as the $-.32$ for Group I.

Although these results purport to throw some light on a most significant matter in psychology, one, however, is restrained from becoming rhapsodical when he scrutinizes the nature of the results which warrant nothing but tentativeness, and discussion of a more or less speculative nature. The data do tend to support the theory that the processes involving the acquisition, association, and retention of facts and symbols atrophy with disuse. Those two classes which had been active in study suffered no age handicap, but a penalty for inactivity was imposed on the group long absent from course work. Thorndike's insight that "adult learning is itself probably a partial preventive or cure for adult inability to learn" receives some experimental validation.

On a speculative basis one might deduce that age as a deterrent to learning is probably not due to physiological changes at all but to the traits, methods, and habits which may be classed under study

habits. The group which exhibited some depreciation with age was not trained in the modes, methods, and even the mild chicaneries that are effective for doing well in objective examinations. Their experiences, which were comparatively remote, included preparation for the essay type of examination which is less severe in the preciseness and comprehensiveness of its requirements. For not knowing definite techniques of study that facilitate success on the objective examination, the uninitiated were penalized. The older studied more but it did not avail them proportionally.

This bespeaks a qualitative deficiency which might be made up by experiences in subsequent courses.

The resuscitation which occurs with active study, besides improving the effectiveness of study methods, may involve the nervous system in a manner analagous to that in which training and exercise improves muscles deteriorated by desuetude. If that be true, it supposes a degeneration of the nervous system. There is no indisputable evidence that such is true. There has been presented evidence which suggests that mental ability can be revived and that being vigorously active in learning is a preservative of mental alertness. There probably is truth from the scientist's standpoint, in the current phrase, "being mentally alive."

But if one becomes "mentally alive" in passing objective examinations, does he become more "mentally alive" in his conversation, in his current reading, and his reactions to political facts and fables? The writer suspects that the measured mental deterioration with age, which is apparent by a measurement of learning, is not a function of mental depreciation or decline of the mental ability curve, but primarily due to a comparative inefficiency in method of study, such as less intensity in study, and an emotional attitude different from that possessed by those who have been made aware of certain needs (natural or otherwise) by experience in previous courses.

In Thorndike's experimentation[1] considerable supremacy over the ages fourteen to eighteen is shown in learning Esperanto by the ages twenty to thirty. It is most improbable that the marked differences represented a differential in learning ability but rather a difference in interest or avidity of attack. There is nothing in the growth curves for mental ability to explain such differences.

The writer suspects that the improvement in "mental alertness" that comes from serious study is specific in its nature. It is probably

[1] Thorndike, E. L.: "Adult Learning." *Loc. cit.*, p. 46.

limited to the means, methods, and materials involved in the pursuit of university credit. Probably the mind does "sharpen" from rigors and severities of course work but does so in learning to adjust itself better by acquiring the techniques and habits for effective study. Holding that the minds of these elementary school teachers become keener for their classroom activities because they were exercised in exacting study, harks back to theories of mental discipline. That their minds become more active and alert by study is not denied, but such improvement is due to the interests and feelings that come from obtaining facts and theories which constitute in the realm of the mental a sort of currency put into circulation. It is behavior with this currency that makes for what one may call mental vigor and alertness. Formal study furnishes people with a pabulum for the mental processes.

A conclusion requires the curtailment of speculation, and it may be stated that there is suggestive evidence that mental ability or learning maintains itself at least for elementary school teachers up to the age of fifty if it is exercised by the severities of course work for credit. The theory supported in this exposition is that disuse, as expressed by absence from the methods used in acquiring facts and information for university study, is the explanation for apparent disability with advancing age. Thus a mode of adaptive behavior requiring certain preciseness and self-discipline needs to be re-acquired by those resuming formal study after considerable absence from it. The groups for whom the extra mural classes of this experiment were preceded by several or many courses showed no age handicap. One may tentatively conclude that at least up to the ages of fifty there is no decline in learning ability with age for those who are actively engaged in study; but probably there is a slight disability, which can be overcome by the resumption of study, for those who have not been engaged in vigorous learning for a period of several or more years.

THE VALIDITY OF TEACHERS' JUDGMENTS OF DIFFICULTY IN CURRICULAR MATERIAL[1]

H. E. SMITH

University of Alberta

In the construction of an educational test it is usually desirable to arrange the test items in an order of increasing difficulty. For this purpose one or both of two methods may be employed. The test may be administered in try-out form to an adequate number of pupils and the per cent of correct responses to each item computed. Or the items may be submitted to a number of judges who may be asked either to rank the items in order of difficulty or to rate them on some arbitrary scale of difficulty. The first method, that of determining the per cent of correct responses, involves a great amount of labor and expense. In any case, before the test is administered at all, it is desirable to have the items arranged as nearly as possible in the true order of difficulty. Hence the validity of judgment is a matter of considerable consequence to the test maker.

It is the purpose of this paper to summarize the results of a study of the validity of judgments of difficulty made by three types of judges in different types of subject matter.

The materials used consisted of 1391 test items taken from the try-out forms I and II of the revised Stanford Achievement Tests. Seven subjects were chosen: arithmetic problems, history, geography, literature, language usage, word meaning, and physiology. The two forms of the seven subjects yielded fourteen "sets" of test items, each set containing approximately one-hundred test items.

The actual difficulty of the items was determined by administering the tests to an approximate total of four hundred pupils in Grades II to IX in five selected American school systems. The per cent of correct responses for each item was computed, the resulting per cents thus constituting inverse criteria of the true difficulty of the items.

Two hundred seventy-nine sets of ratings were obtained from persons classified as follows: one hundred twenty-five experienced teachers, one hundred twenty-five inexperienced teachers, and twenty-nine who may be considered experts in test construction. Experienced teachers were those engaged in teaching in any grade from III to IX, practically all being teachers in California schools. Inexperienced

[1] This article summarizes one phase of the author's Ph.D. thesis written at the University of California under the direction of Dr. G. M. Ruch.

teachers were teachers-in-training who as yet had had no actual classroom teaching experience. Experts in test construction were men well-known in the field of test-making in different parts of the United States.[1] The first group of persons was asked to choose a set of items in the subject in which he or she had the most extended teaching experience. The inexperienced were asked to choose a subject with the content of which he or she was most familiar. The experts in test construction were given their subjects somewhat arbitrarily but an effort was made to have the assigned subjects fall in as closely as possible with the general interests of the person concerned. In almost all cases each judge rated the difficulty of the one hundred or more items in one set only, an exception occurring in the case of the experts, most of whom rated two or more sets.

The sets were sent out in mimeographed form, the items having been arranged in a purely random order of difficulty. The directions for rating as given to the judges were as follows:

> The rating should be done on a scale of 1 to 10. Will you kindly look over the questions, pick out the easiest ones (about one-tenth of the whole number), and place a "1" opposite each. Do the same for the most difficult and put a "10" opposite each of these. After each of the remainder put some number between 1 and 10 to indicate its relative difficulty. Perhaps when you have finished you will have about as many 2's and 3's, etc. as you have 6's or 7's or 8's, etc. Or perhaps you will think that the majority of the questions should fall along the middle ranges of difficulty, as 4's or 5's or 6's. You are to use your own judgment.
>
> The questions were designed by well-known test makers for all the grades from III to VIII, and it was thought that some would be very easy for Grade III people and some rather difficult even for Grade VIII. The questions will ultimately be published in test form and will then be arranged on a difficulty basis.

A ten-point scale was used as representing practically the upper limit of ability of human judgment to differentiate qualities of the sort involved in the task set.

It is apparent that the scale is based upon the relative rather than upon the absolute difficulty of the items. By adopting this point of view two advantages were gained. First, it was possible to secure from the judges the general use of the entire range of the rating scale. In the second place, it was possible to avoid the otherwise inevitable

---

[1] To all those who acted as judges in connection with this investigation the writer owes a debt of gratitude. The debt is especially great to the following, whose time was otherwise so thoroughly occupied: Drs. A. Dvorak, H. A. Greene, Noel Keys, C. W. Odell, Jacob Orleans, W. J. Osburn, H. H. Remmers, G. M. Ruch, Peter L. Spencer, G. D. Stoddard, J. C. Tressler and M. J. Van Wagenen.

confusion in the minds of the judges relative to the grade-status or type of pupil which should be taken as standard.

The two hundred seventy-nine sets of judgment ratings were distributed as shown in Table I.

TABLE I.—THE NUMBER AND DISTRIBUTION OF SETS OF JUDGMENTS USED IN THIS STUDY, ARRANGED ACCORDING TO SUBJECT AND TO TYPE OF JUDGE

Subject	Experienced		Inexperienced		Expert	
	Form I	Form II	Form I	Form II	Form I	Form II
Arithmetic problems	14	14	15	10	..	4
Word meaning	5	5	6	6	..	4
Physiology	3	5	6	6	..	5
History	14	14	12	12	1	3
Geography	14	7	12	8	..	4
Literature	5	9	10	6	..	5
Language usage	9	7	8	8	..	3
Column totals	64	61	69	56	1	28

VALIDITY OF JUDGMENTS

Validity is ordinarily defined as referring to the degree to which a test measures what it is claimed to measure. As here used it refers to the degree of correspondence between a set of obtained ratings of difficulty and the true or empirically determined difficulty of the test items. The assumption has necessarily to be made that a true criterion of difficulty has been secured, but on the basis of this assumption any coefficient of correlation between a set of obtained judgments of difficulty and the criterion of difficulty is properly termed a measure of the validity of the obtained judgments.

Each set of ratings was correlated by the Pearson product-moment method with the appropriate criterion of difficulty. Since it is not feasible to detail here the two hundred seventy-nine results, the coefficients have been averaged by subject and by type of judge and are so presented in Table II. The propriety of averaging correlation coefficients is, of course, open to question, but in view of the narrow range of the measures included within each category it was felt that the mean value represented each group fairly.

Table II reads as follows: For arithmetic problems an average coefficient of correlation of .86 was found between the judgments made by twenty-eight experienced teachers and the criterion of difficulty.

TABLE II.—AVERAGE CORRELATIONS BETWEEN JUDGMENTS OF DIFFICULTY AND ACTUAL DIFFICULTY OF THE TEST ITEMS, ARRANGED BY SUBJECT AND BY TYPE OF JUDGE

Subject	Type of judge							
	Experienced			Inexperienced			Expert	
	N	M	PE_M	N	M	PE_M	N	M
Arithmetic problems	28	.86	± .005	25	.76	± .008	4	.80
Word meaning	10	.82	± .013	12	.70	± .009	4	.77
Physiology	8	.62	± .024	12	.54	± .016	5	.57
History	28	.60	± .009	24	.42	± .012	4	.51
Geography	21	.56	± .016	20	.46	± .012	4	.54
Literature	14	.46	± .014	16	.43	± .012	5	.45
Language usage	16	.45	± .019	16	.41	± .019	3	.45
Totals	125			125			29	
Column means		.62			.53			.58

N = the number of judgment sets averaged.
M = the mean coefficient of correlation.
PE_M = the probable error of the mean.

For twenty-five inexperienced teachers the average coefficient was .76. For the four experts in test construction the average coefficient was .80.

It will be noted that, in the case of the experienced teachers, the coefficients range in descending order from .86 in arithmetic to .45 in language usage. For the inexperienced teachers the order is identical, except in the case of history where a relatively low average appears.

In view of the small number of judges in the expert group it is worthy of note that the same descending order of correlation again occurs, except in the same subject, history.

Another point of interest is that the averages obtained by the experts stand uniformly lower than those secured by the experienced group, but uniformly higher than the averages secured by the inexperienced group.

Only four of the differences between the average coefficients of the experienced and inexperienced groups are statistically significant while the remainder are not. For this determination the formula used was

$$PE_{Diff.} = \sqrt{PE_1^2 + PE_2^2}.$$

The formula as given assumes an absence of correlation between the measures compared—an assumption completely met in the case of independent judgment ratings. The differences, with their respective PE's, and the ratio of differences to $PE_{Diff.}$, are shown in Table III.

TABLE III.—DIFFERENCES AND PE DIFFERENCES BETWEEN THE AVERAGE COEFFICIENTS OF VALIDITY OF THE EXPERIENCED AND THE INEXPERIENCED TEACHERS

Subject	Differences	PE_{Diff}	$\frac{Diff.}{PE_{Diff.}}$
Arithmetic problems	.10	.009	10.3
Word meaning	.12	.016	7.5
Physiology	.08	.029	2.9
History	.17	.015	11.3
Geography	.10	.020	5.0
Literature	.04	.018	2.0
Language usage	.04	.027	1.6

In arithmetic, word meaning, history, and geography, the differences exceed four times their probable errors; in the other three subjects the differences are not statistically significant. In the interpretation of these PE values it is to be noted that several of the comparisons are based upon numbers so small as to render the comparisons of doubtful value. Theoretically, not fewer than twenty-five cases should be included in order to justify the use of the $PE_{Diff.}$ formula.

The average coefficients of validity of judgment presented in Table II are probably higher than students of the recent literature on the reliability of teachers' school marks would have anticipated. They are certainly higher than the validity coefficients reported by previous investigators, Buckingham, Comin, and Orleans and Sealy.

Buckingham submitted to a number of teachers the fifty preferred list words of his spelling scale with the request that they be arranged in their order of difficulty for children to spell. Two hundred arrangements were secured. Of these Buckingham writes: "They differed widely—so widely that whatever may be the value of a consensus of many individuals, the trustworthiness of the judgment of a single teacher appears to be of almost no value."[1] Individual correlations of validity are reported for only five of the arrangements and on the

[1] Buckingham, B. T.: "Spelling Ability." Teachers College, Columbia University, 1913, p. 70.

basis of only ten of the words. The coefficients are of the order 0.12. The combined judgments of the two hundred teachers, when correlated with the true order of difficulty of the words as determined by "trying them out" on pupils, yielded a coefficient of 0.79. This agreement Buckingham regarded as being quite satisfactory.[1]

Comin submitted a list of twenty-three arithmetic problems to twenty experienced teachers who were asked to estimate the per cent of pupils who would solve each problem correctly if given ten minutes for each.[2] The problems were then given to about fifteen hundred pupils in Grades V, VI, VII and VIII in one school. In this way it was possible to determine what problem actually was easiest, and the order of difficulty for the entire list. The only coefficient reported was that between the average rank of the twenty teachers' estimates and the real difficulty rank, a coefficient of 0.74. The impression is left that the validity of the individual teacher's judgment is inconsiderable even in subject matter where judgment ought to be at its best.

The investigation of Orleans and Sealy was identical in procedure with that of Comin but involved the relative difficulty of only five arithmetic problems. The conclusions were likewise of the same general tenor.[3]

It is probable that different factors have operated to produce the very much higher validity coefficients in the present study than in the studies just reviewed. In contrast with the four-grade range of difficulty represented by Comin's arithmetic problems the problems in the Stanford Achievement Tests represent a difficulty range of at least eight grades. The task of judging difficulty in Buckingham's spelling list is probably of the same sort as the task of judging difficulty in language usage. In either case the average adult has probably become so habituated to the correct forms as to have lost any keen appreciation of the relative difficulty of the commoner words and phrases.

Summary

1. In the present study the validity coefficients of judgment ratings obtained from experienced teachers ranged from an average of 0.86 in arithmetic in a descending order through word meaning, physiology,

[1] *Ibid.*, p. 74.

[2] Comin, Robert: Teachers' Estimates of the Abilities of Pupils. *School and Society*, Vol. III, p. 67.

[3] Orleans, J. S. and Sealy, G. A.: "Objective Tests." World Book Company, 1928, p. 27.

history, geography, and literature, to an average coefficient of 0.45 in language usage.

2. The average validity coefficients in arithmetic, word meaning, history, and geography, for experienced teachers, were significantly higher than the corresponding averages for inexperienced teachers.

3. The average validity coefficients for the group of experts in test construction fell, with remarkable uniformity, between the coefficients for the experienced and inexperienced groups.

4. The unit validity coefficients obtained in this study were apparently very much higher than were those found by previous investigators.

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION

CONDUCTED BY FRANCES M. FOSTER

Educational Problems for Psychological Study, by Goodwin Watson and Ralph B. Spence. New York: The Macmillan Co., 1930. Pp. XII + 352.

Although this book is built upon the framework of an earlier one, the additions and modifications are so substantial as to justify the new title. It is an extraordinarily good book, original and stimulating. The plan of the authors has been first to state a number of concrete problems which the teacher or social worker might meet in the course of the day's work. Each problem is followed by a number of questions which, theoretically, can be answered by consulting some one or other of the references given in the extensive lists at the end of each of the chapters. A chapter is composed of a number of these related problems. The main ones deal with such topics as studying, measurement, general method, curricula, special methods, character training, problems of administration, individual differences, emotional conditioning, social attitudes, and group and community relationships. Since the young student may lose himself in the wilderness of titles, an introductory chapter shows him how to get his information economically. These bibliographies are surprisingly full and comprehensive and, in the reviewer's opinion, constitute, with the appendices, the most valuable part of the book.

The authors are catholic in their tastes. Every problem to them is an educational one. There is, however, more emphasis placed on the problems of the social worker than on those of the classroom teacher, and this bias gives a queer slant to the book. The authors do not seem to be quite at home in the classroom; in fact the book reads as if neither of them had ever taught in a regular elementary or secondary school, but had lived their lives interviewing "problem cases." The authors are entitled to their opinion, but the reviewer

believes the book would have been improved if it had had more contacts with classroom teaching (not disciplinary) problems.

Teachers of educational psychology, besides finding it a gold mine of references, will probably find it of greatest use as a source of exercises to supplement a more formal course in their subject. Very few of them, I am confident, will be able to solve the problems any better than their students. They will thus be stimulated to further reading and study. For the teachers of social and abnormal psychology, it will be well-nigh indispensable. And for the teacher in the classroom, not afraid of a little hard digging, it will solve some of the more important problems with which she is confronted.

PETER SANDIFORD.

University of Toronto.

Psychology, Revised Edition, by R. S. Woodworth. New York: Henry Holt and Co., 1929.

The new edition of Professor Woodworth's widely used textbook is a notable revision both in content and organization. Many of the changes of content are the obvious necessities due to the evolution of psychology in the ten years since the first edition was written. "Instinct" is buried, as a term no longer useful in human psychology. Heredity is considered more critically and at greater length. New experimental evidence is introduced throughout the book to enlarge the student's contact with the objective fundamentals of the science.

The organization of the text is the most conspicuously novel feature, and, in the words of the author is "frankly experimental." After an opening chapter introducing the problems of psychology by the examination of the life history of an individual, long chapters on intelligence, memory, learning, and heredity follow in that order. These chapters each treat important topics in a thorough and interesting manner, but each has little relationship to the others. After this beginning, which occupies about half of the book, the author begins the business of outlining a systematic psychology, with a chapter on the stimulus and the motive. In the introduction to the earlier edition, which is quoted in the present one, Woodworth states his task as "to represent the present state of a very active science." From the organization of the revision, the reader may well ask, "What is this

'present state'?" Is it chaos? The treatment of the most important matters of psychology in a form which implies that the order of presentation is an accidental consideration of no importance, almost leads to this impression. In a text in physics it is perhaps of no import whether "sound" or "light" is considered first. Psychology, however, might better present a unified view of its subject-matter, human nature, throughout. This is not adequately done by a series of almost isolated topics, precariously integrated by a chapter on personality at the end.

The new "Psychology" retains and extends a feature which was one of the greatest values of the first edition. This is Professor Woodworth's ability to maintain an equilibrium on a controversial and complex issue, treating it fairly, understandably and interestingly. He admirably avoids the dilemma of indoctrinating the student on one hand, or leaving him in a haze of uncertainty on the other.

LAURANCE F. SHAFFER.

Carnegie Institute of Technology.

Diagnostic and Remedial Work in Arithmetic Fundamentals for Intermediate Grades, by May Lazar. Bureau of Reference, Research and Statistics, Board of Education of the City of New York, 1928. Pp. 203.

"The aim of the investigation was to develop a method of diagnosis of difficulties in the fundamentals, a method of remedial work, and a practical method of using standard tests and practice material as a part of classroom routine." The subjects were Grade VIA pupils, twenty-three boys and twenty-two girls. The median scores on intelligence tests and on tests in arithmetic fundamentals were slightly above the average and grade norms, but the variability of the scores was wide. The low scores on the arithmetic tests were sign posts indicating the need, not of more intensive drill on arithmetic fundamentals, but of specific diagnosis to detect pupil difficulties and to provide remedial treatment. The technique of diagnosis consisted of five methods to detect individual pupil's errors and their causes.

Inventories of poor work habits and of errors (pp. 44–71) are stated specifically and comprehensively. Each is an excellent adaptation of psychology of learning as applied to the fundamental processes in arithmetic. Violations of maxims are noted, such as: . . : "Other

things being equal, form no bond that will have to be broken"; or "A change in the situation means expectation of an altered response"; also, the underpotency of "no" or "not" in a teacher's reply, "No, 5 + 3 are not 9, 5 + 3 are 8."

Remedial instruction is equally as important as are diagnostic methods. Five techniques were developed and were used during ten minutes of the daily arithmetic period. Summaries of remedial measures for each type of pupil difficulty are stated (pp. 96–111) as adequately as the diagnostic inventories.

It is unfortunate that the measurement of class improvement was presented as a series of differences between class medians but with no reference to the statistical significance of those differences. However, the results in the pupils' realization and correction of their deficiencies were most gratifying, also the changed attitudes of pupils and teacher toward each other.

Twelve excellent case studies are included. This investigation should be of distinct value and help to the classroom teacher in directing her toward those techniques for discovering and overcoming individual pupil difficulties during the regular period devoted to the study of arithmetic. The joy of more genuine understanding of pupils' habits of thinking and learning and of the pupils individually awaits such a teacher.

WINONA M. PERRY
Teachers College, University of Nebraska.

Some College Students and Their Problems, by Luella C. Pressey. Columbus: The Ohio State University Press, 1929. Pp. VI + 97.

This is a record of forty case studies illustrating various problems of study, health, family, social relations, morality, and vocational preparation. The writer has drawn from her wide experience in rehabilitation work with college students. The accounts are written simply in non-technical language. The book is primarily for undergraduates.

There is no magic in Mrs. Pressey's methods. We hear nothing of dream analysis or of multiple personality. The author does possess, however, an uncanny ability to get to the root of the problem and to discover what curative measures will be effective. She is never content to stop with a superficial treatment of symptoms.

The reviewer regards this as a valuable introduction to the study of maladjusted students. The book can hardly fail to increase the reader's intelligent sympathy for such persons as are described.

MELVIN RIGG.

Kenyon College.

Practical Statistics for Teachers, by Marion E. Macdonald. The Macmillan Co., New York, 1930. Pp. 176. $1.60.

Greene's "Workbook in Educational Measurement," published during 1928, was followed in 1929 by Lindquist and Stoddard's "Study Manual in Elementary Statistics." The volume here reviewed comprizes the third in this series of useful, practical, and inexpensive publications. Each has its own peculiar merits. Miss Macdonald's would appear to be addressed specifically to teachers who are concerned in statistics primarily as a technical device to be used as part of the routine administration of the classroom. It is not inspiring; it is not explanatory of fundamental principles and procedures; it does not help one to decide which out of an extensive assortment of different techniques would be the best procedure to use under given circumstances. It is not flexible. It is essentially a drill-book in the routine techniques necessary to a working knowledge of elementary educational statistics. And it fulfils this purpose excellently. To the harassed young normal school teacher, concerned with making a good practical job of her teaching rather than with the subtleties of research, it must come as a veritable godsend. Explanations are simple and clear. Contentious issues are sensibly avoided. The working out of problems is facilitated and even made pleasant by the arrangement of material. (Indeed there is here a possible danger that, having had so much done for them in their learning exercises, students will later find themselves unable to work out even simple problems without the aid of the textbook.) The binomial theorem, curve fitting, partial and curvilinear correlations are omitted. Among the appendices are to be found answers to problems, and an excellent set of objective tests on the subject-matter of the work-book.

The inner margins of the pages of the book are perforated, an interesting arrangement, which, however, one cannot evaluate, since it is difficult to comprehend exactly what is to be done with the pages when once they have been torn out. If it is intended that they should

be handed in to the instructor, one might recommend that a few holes punched appropriately near the perforated margin would make it possible to retain the corrected pages in a looseleaf notebook.

All in all, as an elementary text, the book is to be recommended.

O. L. HARVEY.

Graduate School of Education, Harvard University.

PUBLICATIONS RECEIVED

Educational Psychology

Cahill, Beatrice Hunter: *Pupil Guidance.* Boston: The Colonial Press, 1929, pp. XIV + 97.

Collings, Ellsworth and Wilson, Milbourne O.: *Psychology for Teachers.* New York: Charles Scribner's Sons, 1930, pp. XVIII + 474.

Gates, Arthur I.: *Psychology for Students of Education.* New York: The Macmillan Co., 1930, pp. XV + 612.

Laycock, Sam R.: *Adaptability to New Situations.* Baltimore: Warwick and York, Inc., 1929, pp. XI + 170.

Starch, Daniel: *Experiments and Exercises in Educational Psychology.* (Third edition—revised by author and Harry F. Latshaw.) New York: The Macmillan Co., 1930, pp. IX + 254.

General Psychology

Baar, J.: *Psychology.* New York: Globe Book Co., 1928, pp. VII + 126.

Bridges, J. W.: *Psychology Normal and Abnormal.* New York: D. Appleton and Co., 1930, pp. XXII + 552.

Fox, Charles: *Practical Psychology.* New York: Harcourt Brace and Co., 1928, pp. XVII + 180.

Johnson, Wendell: *Because I Stutter.* New York: D. Appleton and Co., 1930, pp. XV + 126.

Pillsbury, W. B.: *The Essentials of Psychology* (third edition). New York: The Macmillan Co., 1930, pp. IX + 466.

Spencer, W. Wylie: *Our Knowledge of Other Minds.* New Haven, Connecticut: Yale University Press, 1930, pp. 145.

Swift, Edgar James: *The Psychology of Childhood.* New York: D. Appleton and Co., 1930, pp. X + 431.

Titchener, Edward Bradford: *Systematic Psychology: Prolegomena.* New York: The Macmillan Co., 1929, pp. XI + 278.

Troland, Leonard T.: *The Principles of Psychophysiology.* New York: D. Van Nostrand Co., Inc., 1929, pp. XIX + 429.

Warren, Howard C. and Carmichael, Leonard: *Elements of Human Psychology.* Boston: Houghton Mifflin Co., 1930, pp. VIII + 462.

Wheeler, Robert F.: *The Psychology of the Spirit.* Boston: Richard G. Badger, 1929, pp. 250.

Educational Sociology

Peters, Charles Clinton: *Foundations of Educational Sociology.* New York: The Macmillan Co., 1930. pp. XV + 476.

Smith, Walter Robinson: *An Introduction to Educational Sociology.* Boston: Houghton Mifflin Co., 1929, pp. XX + 456.

Administration

Angell, Robert Cooley: *A Study in Undergraduate Adjustment.* Chicago: The University of Chicago Press, 1930, pp. IX + 164.

Hertzler, Silas: *The Rise of the Public High School in Connecticut.* Baltimore: Warwick and York, Inc., 1930, pp. XX + 258.

Reeder, Ward G.: *The Fundamentals of Public School Administration.* New York: The Macmillan Co., 1930, pp. XII + 579.

Prepared by the Research Council of the Music Supervisors National Conference in cooperation with the National Bureau for the Advancement of Music: *Survey of College Entrance Credits and College Courses in Music.* New York: National Bureau for the Advancement of Music, 1930, pp. VI + 209.

Scales, Drill, Workbooks and Tests

Adams, Maude and Knight, F. B.: *The Basic Drills in a Minimum Vocabulary of French.* New York: Rand McNally and Co., 1929, pp. 92.

Adams, Maude and Knight, F. B.: *The French Work Book.* New York: Rand McNally and Co., 1929, pp. 90.

Benz, H. E., and Knight, F. B.: *The Efficiency Book for High Schools* (Teacher's Edition). New York: Rand McNally and Co., 1929, pp. IV + 77.

Glenn, Earl R. and Obourn, Ellsworth S.: *Instructional Tests in Physics* (Teacher's Manual). Yonkers-on-Hudson, New York: World Book Co., 1930, pp. 14.

Glenn, Earl R. and Obourn, Ellsworth S.: *Instructional Tests in Physics for High Schools and Colleges.* Yonkers-on-Hudson, New York: World Book Co., 1930, pp. IV + 59.

Glenn, Earl R. and Welton, Louis E.: *Instructional Tests in Chemistry* (Teacher's Manual). Yonkers-on-Hudson, New York: World Book Co., 1929, pp. 13.

GLENN, EARL R. and WELTON, LOUIS E.: *Instructional Tests in Chemistry.* Yonkers-on-Hudson, New York: World Book Co., 1930, pp. IV + 76.

KNIGHT, F. B. and MCCLURE, W. E.: *Arithmetic Neatness Scale.* Chicago: Scott, Foresman and Co., 1927,

LANE, ROBERT HILL: *A Work Book for Principals and Supervisors.* New York: The Macmillan Co., 1930, pp. VIII + 263.

LINDQUIST, EVERETT, F. and STODDARD, GEORGE D.: *Study Manual in Elementary Statistics.* New York: Longmans, Green and Co., 1929, pp. VII + 109.

MYERS, G. W. (Editor): *Manual of Directions for Compass Diagnostic Tests in Arithmetic.* Chicago: Scott, Foresman and Co., 1925, pp. 64.

RUGG, HAROLD and MENDENHALL, JAMES E.: *Pupil's Workbook* (to accompany *An Introduction to American Civilization*). Boston: Ginn and Co., 1929, pp. XIV + 80.

RUGG, HAROLD and MENDENHALL, JAMES E.: *Pupil's Workbook* (to accompany *Changing Civilizations in the Modern World*). Boston: Ginn and Co., 1930, pp. XVI + 99.

SANGREN, PAUL V.: *Information Tests for Young Children.* Yonkers-on-Hudson, New York: World Book Co., 1930, pp. III + 28.

WEBSTER, EDWARD HARLAN: *Oral Tests for Correct English.* Yonkers-on-Hudson, New York: World Book Co., 1930, pp. IV + 88.

WELLS, F. L. and KNIGHT, F. B.: *Basic Spelling Units* (for Junior High School Students). New York: Rand McNally and Co., 1929, pp. 104.

WELTE, H. D. and KNIGHT, F. B.: *Standard Service Geometry Work-Book* (Teacher's Edition). Chicago: Scott, Foresman and Co., 1929, pp. XVII + 74.

MONOGRAPHS

EDGERTON, HAROLD A.: *Academic Progress.* Columbus, Ohio: The Ohio State University Press, 1929, pp. 150.

HURD, ARCHER WILLIS: *Problems of Science Teaching at the College Level,* Volume I. Minneapolis: The University of Minnesota Press, 1929, pp. XVIII + 195.

JOHNSON, PALMER O.: *Curricular Problems in Science at the College Level,* Volume II. Minneapolis: The University of Minnesota Press, 1930, pp. XVI + 188.

McCarthy, Dorothea A.: *The Language Development of the Preschool Child.* Minneapolis: The University of Minnesota Press, 1930, pp. XIII + 174.

Tidwell, C. J.: *State Control of Textbooks.* New York: Bureau of Publications, Teachers College, Columbia University, 1928, pp. 78.

General

Collins, Joseph: *Insomnia—How to Combat It.* New York: D. Appleton and Co., 1930, pp. VI + 131.

Garcon, Maurice and Vinchon, Jean: *The Devil.* New York: E. P. Dutton and Co., 1930, pp. 288.

Hocking, William Ernest: *Human Nature and Its Remaking.* New Haven, Connecticut: Yale University Press, 1929, pp. XXVI + 496.

Johnson, Charles S: *The Negro in American Civilization.* New York: Henry Holt and Co., Inc., 1930, pp. XIV + 538.

Lindworsky, Johann: *The Training of the Will.* Milwaukee: The Bruce Publishing Co., 1929, pp. 226.

Lorimer, Frank: *The Growth of Reason.* New York: Harcourt, Brace and Co., 1929, pp. XII + 231.

Seligman, Edward R. A. (Editor-in-Chief): *Encyclopedia of the Social Sciences,* Volume I (Aarenson-Allegiance). New York: The Macmillan Co., 1930, pp. XXVII + 646.

Miscellaneous

Benedict, Agnes E.: *Children at the Crossroads.* New York: The Commonwealth Fund Division of Publications, 1930, pp. 238.

Buchanan, J. Arthur: *Nutriology.* Boston: Richard G. Badger, 1930, pp. 149.

Durea, Mervin A. (Editor): *Subject Index Volume III of Child Development Abstracts and Bibliography.* Washington, D. C.: Issued by Committee on Child Development National Research Council, 1929, pp. 63.

Ewing, Irene R.: *Lipreading.* Manchester, England: Manchester University Press, 1930, pp. IX + 74.

Nash, Jay B. (Editor): *Symposium on Physical Education and Health.* New York: New York University Press Book Store, 1930, pp. 320.

Foreign Publications

Alves, Isaias: *Teste Individual de Intelligencia.* Bahia: Officinas Graphicas d'a Luva, 1928, pp. XI + 175.

Alves, Isaias: *Os Testes e a Reorganização Escolar.* Bahia: A Nova Graphica, 1930, pp. XVI + 255.

Grzegorzewska, Dr. Marja: *Psychologja Niewidomych,* Vol. I. Warsaw, 1930, pp. 348.

Grzegorzewska, Dr. Marja: *Psychologie des Aveugles.* 1930, pp. 32.

Kern, Dr. Benno: *Wirkungsformen Der Übung.* Munster i. Westf: Helios-Verlag G. M. B. H., 1930, pp. IX + 508.

Newhall, Almer M.: *Comments on the Political and Economic Status of the Philippines.* San Francisco Chamber of Commerce, pp. 30.

Pièron, Henri: *Le Développement Mental et L'Intelligence.* Paris: Librairie Félix Alcan, 1929, pp. XII + 95.

Ranschburg, Dr. Paul: *Psychologische Studien.* Budapest: Koniglich Ungarische Universitäts-Druckerei, 1929, pp. VIII + 493.

Elementary Textbooks

Clark, Otis, Hatton: *Modern-school Arithmetic, Book Three.* Yonkers-on-Hudson, New York: World Book Co., 1930, pp. XIII + 402.

Cochran and Eddy: *Pierrille.* Chicago: University of Chicago Press, 1930, pp. XIII + 200.

Evans, Lawton B.: *The Pathfinder.* New York: The Macmillan Co., 1930, pp. XII + 515.

Gates and Huber: *The Work-play Books, Peter and Peggy* (Primer). New York; The Macmillan Co., 1930, pp. VII + 126.

Gates and Huber: *The Work-play Books, Round the Year* (First Reader). New York: The Macmillan Co., 1930, pp. VIII + 167.

Gates and Huber: *The Work-play Books, Friendly Stories* (Second Reader). New York: The Macmillan Co., 1930, pp. XIII + 226.

Gates and Huber: *The Work-play Books, Make and Make-believe* (Third Reader). New York: The Macmillan Co., 1930, pp. X + 310.

Lacey, Ida Belle: *Light Then and Now.* New York: The Macmillan Co., 1930, pp. VII + 209.

Rugg, Harold: *An Introduction to American Civilization.* Boston: Ginn and Co., 1929, pp. XIII + 610.

RUGG, HAROLD: *Changing Civilizations in the Modern World.* Boston: Ginn and Co., 1930, pp. XVII + 633.

WEBSTER, EDWARD HARLAN: *Daily Drills for Better English.* Yonkers-on-Hudson, New York: World Book Co., 1930, pp. X + 388.

CIVIC EDUCATION

DIEMER and MULLEN: *Pupil Citizenship.* Yonkers-on-Hudson, New York: World Book Co., 1930, pp. XI + 339.

PETERS, CHARLES CLINTON: *Objectives and Procedures in Civic Education.* New York: Longmans, Green and Co., 1930, pp. VII + 302.

TEACHING AND LEARNING

ADAMS, FAY and BROWN, WALKER: *Teaching the Bright Pupil.* New York: Henry Holt and Co., 1930, pp. XIV + 249.

LATON, ANITA DUNCAN: *The Psychology of Learning Applied to Health Education through Biology.* New York City: Bureau of Publications, Teachers College, Columbia University, 1929, pp. VI + 103.

MICHELL, ELENE: *Teaching Values in New-type History Tests.* Yonkers-on-Hudson, New York: World Book Co., 1930, pp. IX + 179.

MUSE, MAUDE BLANCHE: *An Introduction to Efficient Study Habits.* Philadelphia and London: W. B. Saunders Company, 1929, pp. 110.

REPORTS OF PROCEEDINGS

Proceedings of the Association of History Teachers of the Middle States and Maryland (No. 27), Montpelier, Vt.: Capital City Press, 1929, pp. 148.

AMBROSE, L. SUHRIE, (Editor): Proceedings of the 1929 Spring Conference Conducted by the Normal-school and Teachers-college Section of the New York Society for the Experimental study of Education: *Problems in Teacher-training* (Volume IV). New York University Press Book Store, Washington Square East, New York City, pp. XIV + 355.